DOCUMENT RESUME

ED 239 560 HE 016 984

AUTHOR	Mentkowski, Marcia; Doherty, Austin
TITLE	Validating Assessment Techniques in an Outcome-Centered Liberal Arts Curriculum: Integrated Competence Seminar. Final Report, Research Report Number Four.
INSTITUTION	Alverno Coll., Milwaukee, Wis.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
PUB DATE	82
GRANT	NIE-G-77-0058
NOTE	69p.; For related documents, see HE 016 980-990.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	College Curriculum; *Competency Based Education; Evaluation Methods; Group Discussion; Higher Education; Liberal Arts; Outcomes of Education; *Problem Solving; *Research Methodology; *Seminars; Speech Skills; *Student Evaluation; *Validity
IDENTIFIERS	*Alverno College WI

ABSTRACT

The Integrated Competence Seminar (ICS), An Alverno College assessment technique used in the outcome-centered liberal arts curriculum, is discussed. The ICS technique allows students to demonstrate integration and transfer of learning in three situations: oral presentation, group discussion, and the In-Basket exercise (decision-making problems). Assessors observe and evaluate performance criteria, and provide students with feedback. In the 1970s, the In-Basket exercise was the most difficult and the most valid in terms of correlation with measures of students' cognitive development and other generic ability. Oral presentation showed mixed results, and the group discussion correlated with other measures in opposite to the expected directions. When age, background, and college program were controlled, there were no significant relationships between the three ICS tasks and other college performance measures. Generally, it was found that the In-Basket exercise most accurately measured abilities of analysis and problem-solving. The results point to the importance of nontraditional assessment techniques, such as In-Basket, and to revision of the measure with attention to the links between group discussion criteria and the social interaction ability it represents. (Author/SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED239560

VALIDATING ASSESSMENT TECHNIQUES IN AN OUTCOME–CENTERED LIBERAL ARTS CURRICULUM: INTEGRATED COMPETENCE SEMINAR

Assessment Committee/Office of Research & Evaluation
ALVERNO COLLEGE

FINAL REPORT TO THE NATIONAL INSTITUTE OF EDUCATION:
RESEARCH REPORT NUMBER FOUR

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Funded by a grant from the National Institute of Education:
Careering After College: Establishing the Validity of Abilities Learned in College for Later Success
(NIE–G–77–0058)

Principal Investigators:
Marcia Mentkowski
Austin Doherty
Alverno College
3401 South 39th Street
Milwaukee, Wisconsin 53215

NIE 016 924

An overview and rationale for our approach to the study of college outcomes, and a summary of the results from the following series of ten research reports, are found in:

Marcia Mentkowski and Austin Doherty. Careering After College: Establishing the Validity of Abilities Learned in College for Later Careering and Professional Performance. Final Report to the National Institute of Education: Overview and Summary. Milwaukee, WI: Alverno Productions, 1983.

Research Reports:

One: Friedman, M., Mentkowski, M., Earley, M., Loacker, G., & Diez, M. Validating Assessment Techniques in an Outcome-Centered Liberal Arts Curriculum: Valuing and Communications Generic Instrument, 1980.

Two: Friedman, M., Mentkowski, M., Deutsch, B., Shovar, M.N., & Allen, Z. Validating Assessment Techniques in an Outcome-Centered Liberal Arts Curriculum: Social Interaction Generic Instrument, 1982.

Three: Assessment Committee/Office of Research and Evaluation. Validating Assessment Techniques in an Outcome-Centered Liberal Arts Curriculum: Insights From the Evaluation and Revision Process, 1980.

Four: Assessment Committee/Office of Research and Evaluation. Validating Assessment Techniques in an Outcome-Centered Liberal Arts Curriculum: Integrated Competence Seminar, 1982.

Five: Assessment Committee/Office of Research and Evaluation. Validating Assessment Techniques in an Outcome-Centered Liberal Arts Curriculum: Six Performance Characteristics Rating, 1983.

Six: Mentkowski, M., & Strait, M. A Longitudinal Study of Student Change in Cognitive Development and Generic Abilities in an Outcome-Centered Liberal Arts Curriculum, 1983.

Seven: Much, N., & Mentkowski, M. Student Perspectives on Liberal Learning at Alverno College: Justifying Learning as Relevant to Performance in Personal and Professional Roles, 1982.

Eight: Mentkowski, M., Much, N., & Giencke-Holl, L. Careering After College: Perspectives on Lifelong Learning and Career Development, 1983.

Nine: Mentkowski, M., DeBack, V., Bishop, J., Allen, Z., & Blanton, B. Developing a Professional Competence Model for Nursing Education, 1980.

Ten: Mentkowski, M., O'Brien, K., McEachern, W., & Fowler, D. Developing a Professional Competence Model for Management Education, 1982.

© Copyright 1982. Alverno College Productions, Milwaukee, Wisconsin. All rights reserved under U.S., International and Universal Copyright Conventions. Reproduction in part or whole by any method is prohibited by law.

ABSTRACT

The Integrated Competence Seminar assessment technique allows students to demonstrate integration and transfer of learning in three situations: Oral Presentation, In-Basket Exercise and Group Discussion. Assessors observe and evaluate performance against specified criteria, and give feedback to students on their performance. A behavioral criteria checklist permits evaluation of inter-rater reliability, and validation of the technique through comparison of quantified assessor judgments with other student performance measures and a battery of external criterion measures administered to students in a longitudinal study of college outcomes (Mentkowski & Strait, 1983).

Assessor judgments correlated in the 70's. The In-Basket Exercise was the most difficult and the most valid in terms of correlation with measures of students' cognitive development and other generic abilities. Oral Presentation showed mixed results, and the Group Discussion correlated with other measures in opposite to the expected directions. When age, background and college program are controlled, there were no significant relationships between the three ICS tasks and other college performance measures, namely, number of credits accumulated and number of competence level units achieved. Thus, the In-Basket had some performance validity, the Oral Presentation is equivocal, and the Group Discussion had relationships opposite to our expections. The finding on the Group Discussion supports our earlier findings with respect to a Social Interaction generic instrument.

Generally, the effort revealed that the In-Basket exercise most accurately measured abilities of Analysis and Problem Solving. The Group Discussion, a measure of Social Interaction, worked less well. The study points to the importance of continuing to develop nontraditional assessment techniques like In-Basket, and to revise the measure with particular attention to the links between Group Discussion criteria and the Social Interaction ability it represents. The Integrated Competence Seminar has since undergone extensive revisions by a group of faculty specializing in assessment design, based in part on the findings of this study.

ACKNOWLEDGEMENTS

We wish to acknowledge the contributions of the following persons:

Office of Research & Evaluation:
Miriam Friedman
Marcia Mentkowski
Eunice Monroe
Michael J. Strait

Assessment Committee (1977–82):
Zita Allen
Jean Birkey
Austin Doherty
Margaret Earley
Miriam Friedman
Georgia Gurria
Pat Jensen
Joan Koehler
Cathleen Krzyminski
Dimitri Lazo
Georgine Loacker
William McEachern
Marcia Mentkowski
Agnes Meysenburg
Marlene Neises
James Newton
Kathleen O'Brien
James Roth
Jean Schafer
Judeen Schulte
Kyle Stewart
Michael Strait
Alice Theine
Christine Trimberger
Allen Wutzdorff

External Assessors:
Robert Scorgie
Jeanne Jung

TABLE OF CONTEXTS

	Page
The Integrated Competence Seminar: A Diagnostic Assessment	1
The Integrated Competence Seminar: An External Criterion Measure	3
Developing a Behavioral Criteria Checklist	3
Establishing Checklist Reliability	4
Analysis of Student Performance Using Quantified Student Scores	4
Consistency of Students' Performance Across Exercise Criteria	12
Validating the Integrated Competence Seminar in Relation to Other Student Performance Measures in the Longitudinal Study	26
Analysis of Integrated Competence Seminar Scores for the Weekday College Longitudinal Sample	32
Relationship between Integrated Competence Seminar and Student and Program Characteristics	37
Relationship between Integrated Competence Seminar and Other College Performance Measures	39
Relationship between Integrated Competence Seminar and Human Potential Measures	42
Summary and Discussion	51
References	54
Appendix I: Behavioral Cirteria Checklist	56

VALIDATING ASSESSMENT TECHNIQUES IN AN OUTCOME-CENTERED LIBERAL ARTS CURRICULUM: INTEGRATED COMPETENCE SEMINAR

Assessment Committee/Office of Research & Evaluation

The Integrated Competence Seminar: A Diagnostic Assessment

The Integrated Competence Seminar, an Alverno College assessment technique, provides an opportunity for a student from any discipline to demonstrate integration of the abilities she has developed at the midpoint of her college education. She also shows how she transfers these abilities to a new and complex simulated situation she is likely to face as a college-educated person.

Prior to completing the Integrated Competence Seminar (ICS), the student has more than likely been credentialed on level 4 of each competence. The student is not credentialed on her performance on the ICS; the assessment measures generalization of her learning to a new situation without the pressure to perform for validation. Thus, the ICS can also be viewed as an external criterion measure of a student's ability to transfer her learning after completing the general education sequence.

This cross-disciplinary simulation assesses students' integration of level 4 of four competences: Communications, Analysis, Problem Solving, and Social Interaction. The Integrated Competence Seminar involves the student in various activities as a member of a decision-making board of citizens who must make a selection among three proposals for distribution of funds for civic improvement. The student takes on the role of a representative of one of six community agencies: health, culture, recreation, education, vocational training or environmental affairs.

In the ICS, the student engages in a series of three exercises over a

four-hour period. First she makes an Oral Presentation. After she studies a set of background information, she delivers a speech to persuade the decision-making board to accept her proposal. Next, she completes the In-Basket exercise, which derives its name from the in-basket on the manager's desk containing letters, memos, minutes and reports, all awaiting action. She solves problems, sets priorities, analyzes, organizes, and makes decisions on seven different items as if she were a board member encountering such situations in her office. In the Group Discussion, the student meets with other "board members" to discuss and make decisions on the final proposal--a common plan for the expenditure of funds.

Off-campus professionals, rather than Alverno faculty, serve as assessors. When off-campus assessors observe students' performance the value of out-of-class learning and assessment experiences are legitimated. Students begin to realize that abilities learned in college must at some point be transferred to off-campus work experiences, and will be judged by others out of their professional perspectives. The assessors thus give an added professional dimension to student assessments.

Assessors observe and evaluate each student's performance against specified criteria (oral and group sessions are videotaped). Each first records his or her behavioral observations pertinent to the criteria and then meets with another assessor to achieve consensus and develop a written evaluative summary. Assessors then schedule individual feedback sessions during which an assessor shares each student's evaluative summary with her, discussing strengths and weaknesses in her performance on the ICS. Consequently, the ICS provides diagnostic information for both students and faculty.

The Integrated Competence Seminar: An External Criterion Measure

The major purpose of the Integrated Competence Seminar is, therefore, diagnostic; it assists faculty to plan the upcoming professional sequence. The measure can also be used as an external criterion measure for curriculum validation studies. Prior to using the ICS as an external criterion measure, however, quantification of the qualitative behavioral observations and evaluative summary is necessary. Such quantification facilitates statistical comparisons of student performance integration across the competences assessed at level 4 although the students may have learned their abilities in different ways and courses. The ICS total score can then also be compared to other external criterion measures to provide information about the strength of the relationships among various student abilities.

Developing a Behavioral Criteria Checklist

In order to translate the assessors' qualitative behavioral observations and evaluative summary into quantifiable data it was decided to construct a list of behavioral criteria which could be used as a checklist for each of the three ICS exercises: Oral Presentation, In-Basket and Group Discussion. In Fall 1978, 29 student ICS files were reviewed. A complete list of behaviors recorded by the assessors was generated. There were no behaviors recorded by the assessors that were not clearly similar to the instrument's behavioral criteria. However, 25% of the behavioral criteria were recorded less than three times across the 29 files. These behaviors were eliminated. The remaining 75% comprised the criteria for the checklists. The question as to whether assessors failed to identify 25% of the specified behaviors because the stimulus did not elicit them, the students did not demonstrate them, or assessors were perceptually unaware of them, awaits further investigation.

A behavioral criteria checklist was thus designed for each of the three exercises: Oral Presentation, In-Basket, and Group Discussion. Assessors respond to each of the criteria by checking the frequency of occurrence of student response along the following scale: never, at least once, frequently, as often as possible. This procedure provides a score for each criterion, an exercise score for each of the three exercises, and a combined exercise score describing overall performance on the ICS. The latter score is called "ICS total score" (Appendix I).

Establishing Checklist Reliability

Two independent evaluators then generated quantitative data from the 29 evaluative summaries using the checklist. Inter-rater reliability reached acceptable levels for each exercise: Oral Presentation, r=.72; In-Basket, r=.75; Group Discussion, r=.79 (see Table 1). These results supported the reliability of the checklist as a quantitative measure of qualitative student performance data from the ICS and demonstrated that assessors tend to record behavioral observations in a reliable manner, allowing independent evaluators to interpret the frequency of the behavior's occurrence on a quantitative checklist.

Once the checklist was found to be a reliable measure, it was instituted in the regular administration of the Integrated Competence Seminar (Fall 1978). This meant that the assessors completed the qualitative behavioral observations and evaluative summary as well as the quantitative checklist from direct observation of the students' ICS performance.

Analysis of Student Performance Using Quantified Student Scores

Another sample of students was then selected for analysis of student performance using quantified student scores obtained from the behavioral

TABLE 1

Inter-rater Reliability of the Integrated Competence Seminar Quantitative Behavioral Criteria Checklist Using the Qualitative Evaluative Summary as Information Base (n=29)

Group Discussion[a]		In-Basket[a]		Oral Presentation[a]		Combined ICS Total Score[b]	
Rater 1	Rater 2	Rater 1	Rater 2	Rater 1	Rater 2	Rater 1	Rater 2
86	88	51	53	82	90	22	23
100	100	95	100	97	96	29	29
83	71	70	69	85	84	24	22
86	88	47	37	67	93	20	21
95	100	74	41	91	90	26	23
83	66	79	62	67	54	23	18
93	95	86	58	97	60	28	21
71	80	42	58	79	78	19	22
12	52	40	39	33	30	8	12
2	35	70	51	30	42	10	13
52	83	40	83	48	69	14	23
93	100	98	88	58	69	25	26
88	97	79	88	73	84	24	27
93	83	40	46	39	60	12	19
69	64	47	51	48	69	16	18
88	71	81	76	55	60	22	21
100	100	81	72	94	75	27	25
81	71	70	39	70	51	22	16
76	97	100	97	73	90	25	28
76	90	67	74	88	90	23	25
98	100	98	100	97	90	27	29
71	54	77	62	82	78	23	19
36	64	60	41	42	57	14	16
95	90	86	83	82	97	26	23
88	78	88	93	94	93	27	26
100	100	91	88	97	78	29	27
40	47	40	32	45	66	12	14
98	100	100	100	100	100	30	30
43	52	47	46	97	100	19	20
r = .79		r = .75		r = .72		r = .78	

[a]Exercise total score in percent assigned to each student on the Integrated Competence Seminar.

[b]Raw score

criteria checklists. Forty-eight students who had completed the Integrated Competence Seminar by Fall 1979 and who were part of a longitudinal sample of students (entering Alverno in Fall 1977) already participating in a study of college outcomes, comprised the sample.

Students completed the Integrated Competence Seminar, and assessors then observed student performance, recorded the demonstrated behaviors, came to consensus on the Group Discussion, evaluated each student against the behavioral criteria and created an evaluative summary. This time, the assessors also completed the behavioral criteria checklists.

The inter-rater reliability of one checklist was first computed. Since Group Discussion was the only exercise where two or more assessors were reviewing student performance, we investigated inter-rater reliability for this exercise only. We correlated the checklist rating of each individual assessor on Group Discussion with the consensus rating (r=.85). Thus, assessors consistently agreed with consensus on an overall low performance or overall high performance rating for each student.

Each student was first assigned an exercise score. The exercise score was derived by adding the student's points across all criteria per exercise (as often as possible = 3 points, frequently = 2 points, at least once = 1 point, not at all = 0 points), and then dividing the student's total by the maximum number of points possible for the exercise.

This yielded a percentage score. For example, the Group Discussion exercise consists of 13 criteria. The maximum number of points is 39 (13 X 3 = 39). If the student's total number of points for the Group Discussion is 20, then her percentage score on the exercise is 51%.

Each student was then assigned an ICS total score. This score, a combination of the three exercise scores, was derived by adding the number

of points from each of the exercises, and dividing the total by the maximum number of points possible across the three exercises.

The distribution of behavioral criteria percent scores per exercise and the ICS total score are presented in Figures 1, 2, 3 and 4.

The graphic representation of the score distributions indicate a bi-modal distribution in the Group Discussion exercise (Figure 1). There were few students with low scores and few students with very high scores. However, approximately 58% of the students obtained a score between 50-80%.

In the In-Basket (Figure 2) only 38% of the students obtained a score between 50-80% and none obtained a score higher than 80%.

In the Oral Presentation (Figure 3) 47% of the students obtained a score between 50-80%.

The graphs clearly show that the Group Discussion exercise was the easiest, and the In-Basket the more difficult one; the Oral Presentation fell between the other two.

Figure 4 shows the total score for all three exercises combined; 78% obtained a score between 50-80%, which indicates that some students who did poorly on one exercise did well on the others.

Total student performance scores on each of the exercises were then compared via paired t-tests. Students performed significantly better on the Group Discussion ($\bar{X}$ = 23.18; SD = 8.68) than they did on the In-Basket ($\bar{X}$ = 17.87; SD = 8.02) (t = 3.3; p <.01). The difference in performance on the Group Discussion and the Oral Presentation was not significant ($\bar{X}$ = 20.79; SD = 6.68) t = 1.96; n.s.); the difference in performance on the In-Basket and the Oral Presentation was also not significant (t = 1.94; n.s.). Clearly, students are more likely to perform better in the Group Discussion than they are on the In-Basket or on the Oral Presentation.

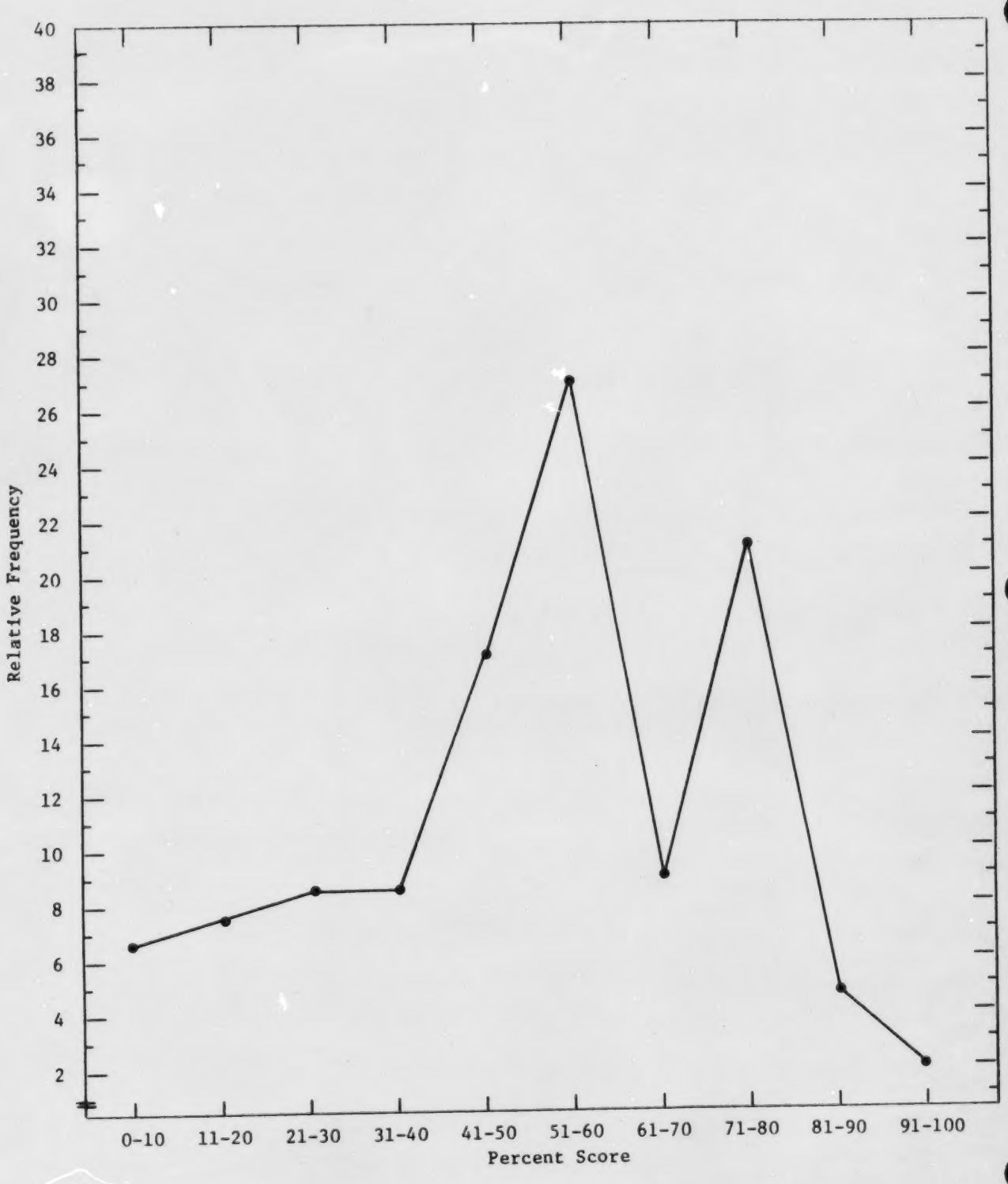


Figure 1. Relative Frequency Distribution of Students' Scores on the Group Discussion Exercise

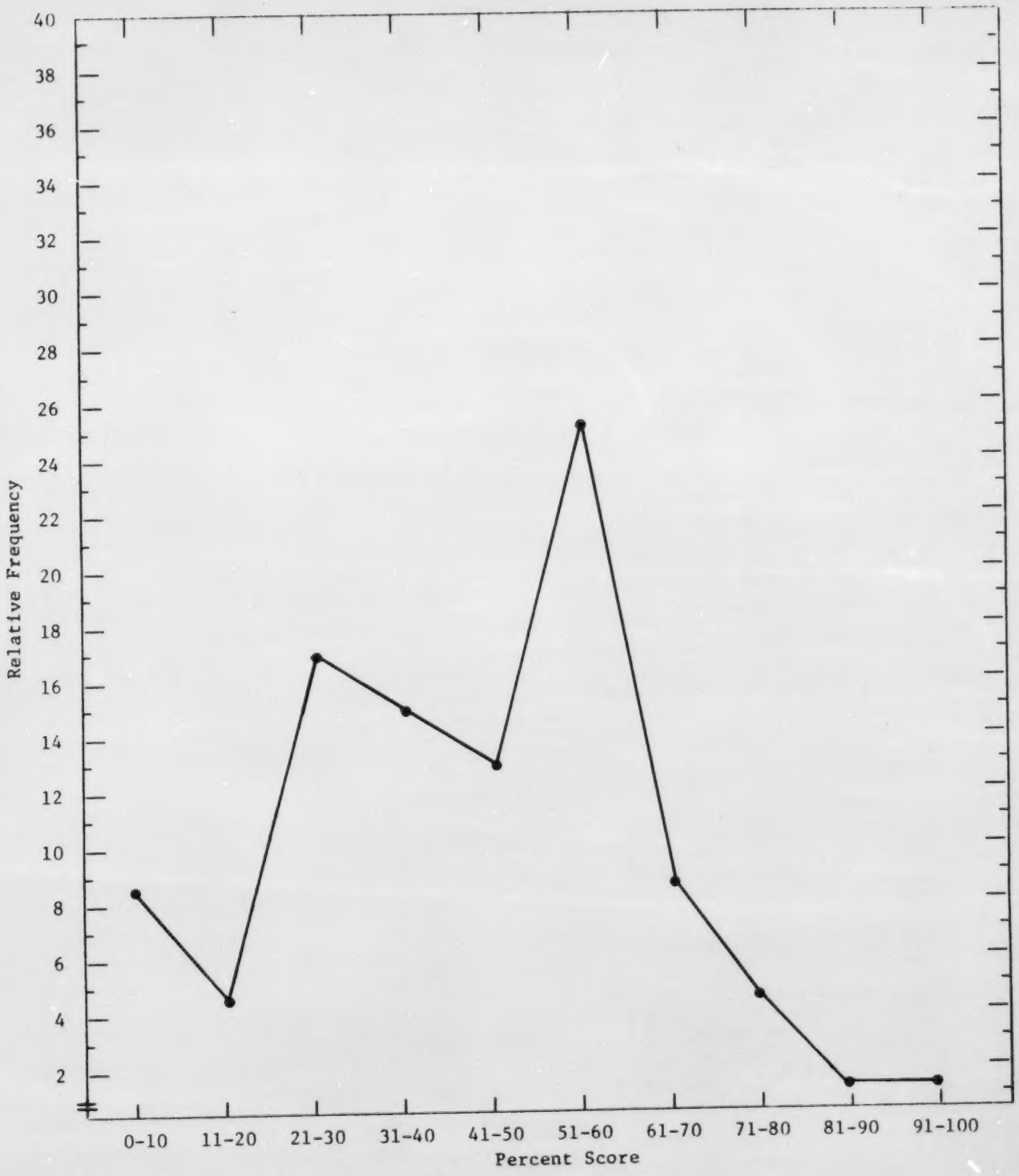


Figure 2. Relative Frequency Distribution of Students' Scores on the In-Basket Exercise

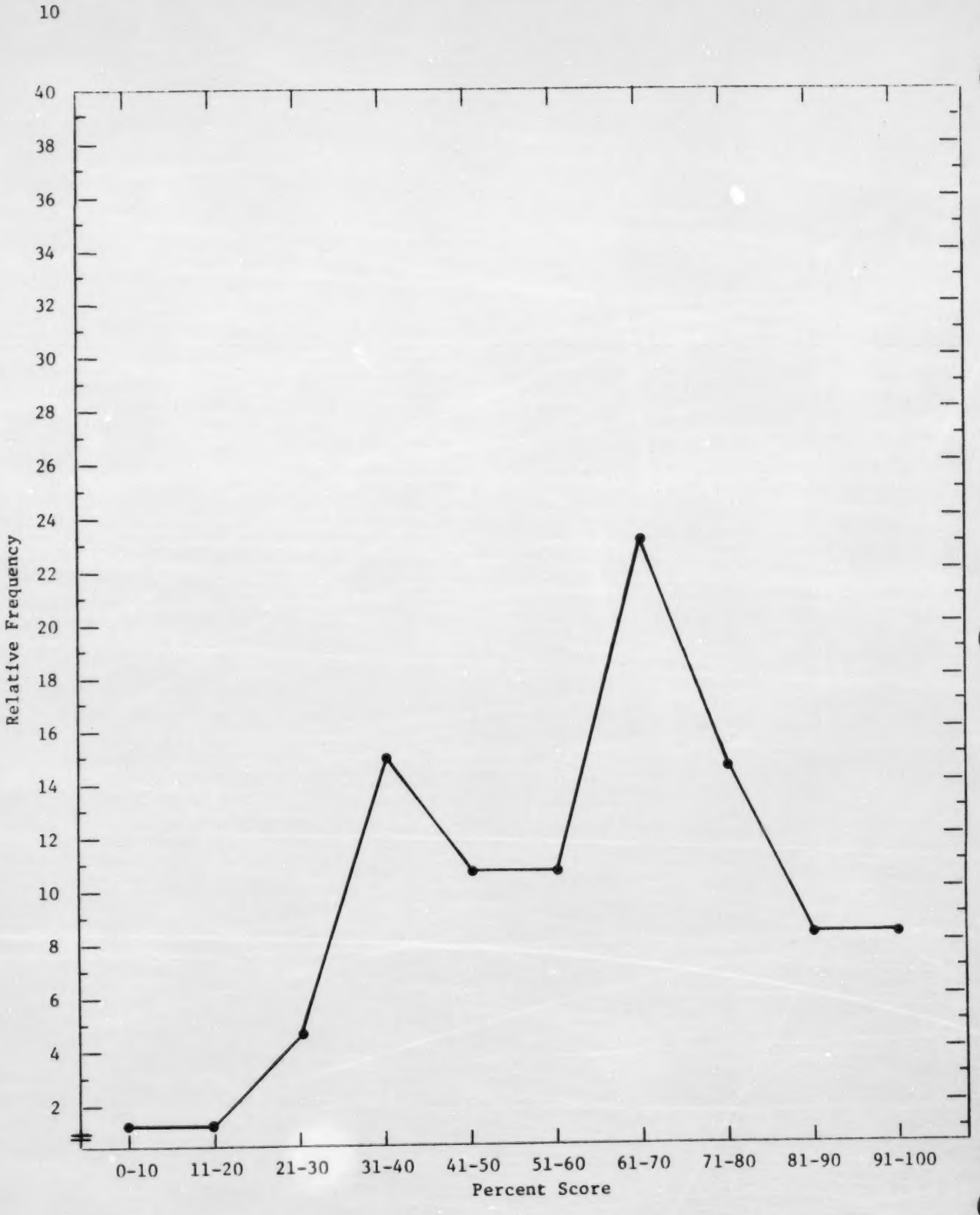


Figure 3. Relative Frequency Distribution of Students' Scores on the Oral Presentation Exercise

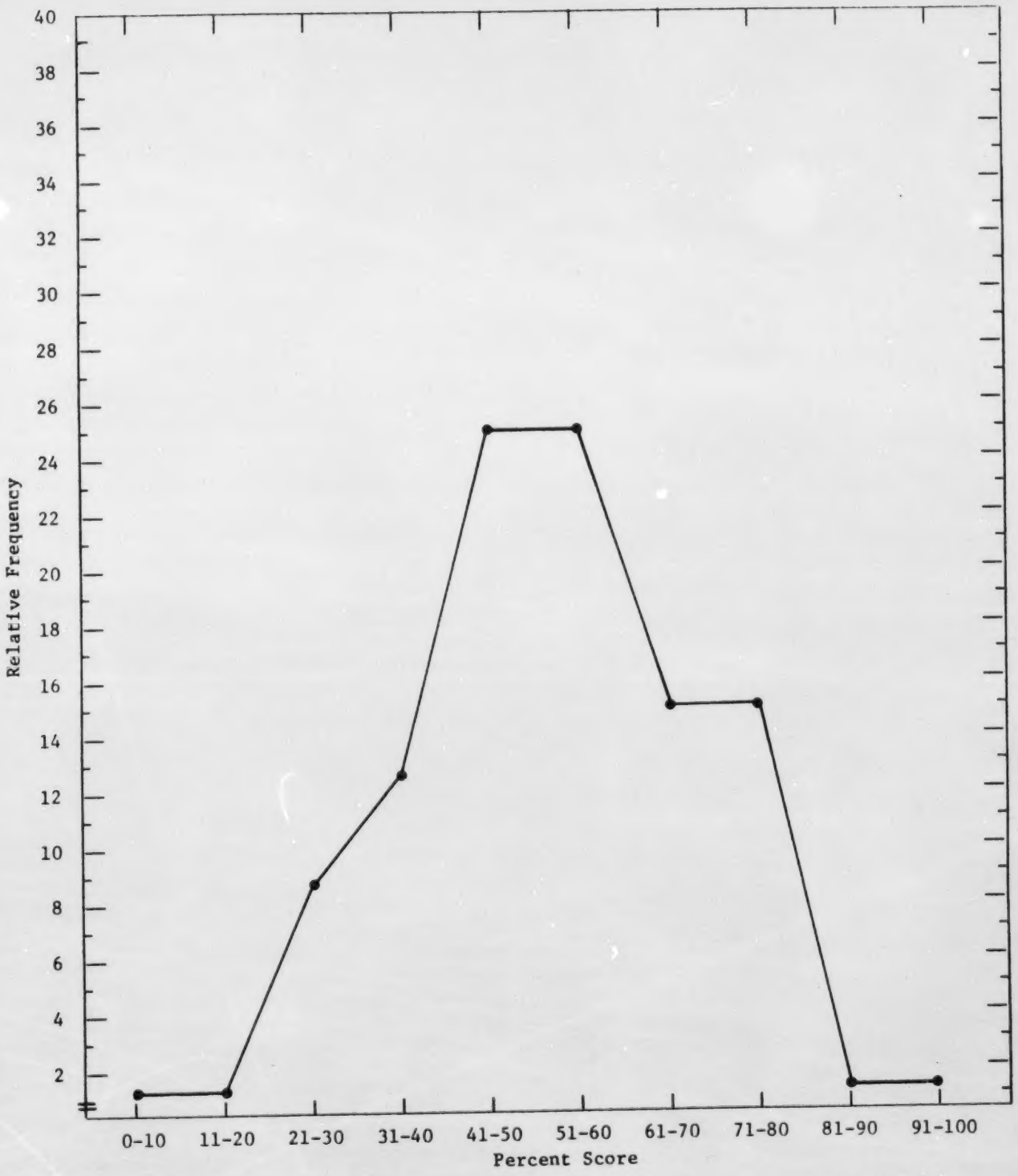


Figure 4. Relative Frequency Distribution of Students' Scores on the Integrated Competence Seminar

Consistency of Students' Performance
Across Exercise Criteria

To what extent did students perform consistently across exercises on the same behavioral criteria?

Students' performance on each criterion was compared to examine the pattern of differences among the behavioral criteria for each exercise. Criteria are grouped under five broader categories: Initiation, Influence, Response, Problem Analysis and Decision-Making. Criteria were not identical across each exercise, of course. Group Discussion has 13 criteria, the In-Basket has 12, and the Oral Presentation has 11. Each student's scale point (0, 1, 2 or 3) per criterion was entered into a comparison with a similar criterion if at least two exercises had a criterion in common.

Table 2 shows the means and standard deviations for the behavioral criteria for each exercise. Several t-tests for paired comparisons were performed, and are entered in Table 3. As shown in Table 3, students demonstrated significantly higher performance on Good Impression in the Group Discussion as compared to In-Basket. When In-Basket criteria were compared to Oral Presentation, students performed consistently higher in the Oral Presentation on Good Impression, Attention and Respect, Confidence, Adaptability, Reading Comprehension and Concern for Details. As for Oral Presentation vs. the Group Discussion, students performed consistently higher in the Oral Presentation on Confidence, Adaptability, Organization and Decisiveness.

Table 4 shows the correlation coefficients between the paired comparisons. Significant positive correlations were obtained on Good Impression, Attention and Respect, and Confidence for the Oral Presentation vs. Group Discussion comparison. Similarly significant positive correlations were obtained on Good Impression and Decisiveness in the In-Basket vs. Oral Presentation.

These results indicate that independent measures of the same behaviors

TABLE 2

Means ($\bar{X}$) and Standard Deviations (SD)
Per Criterion for Each Exercise

	Group Discussion		In-Basket		Oral Presentation	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
INITIATION						
Oral Expression	--	--	--	--	2.02	.65
Clear Writing	--	--	1.70	.98	--	--
Initiative	1.71	.80	1.52	.92		
INFLUENCE						
Persistence	1.67	.91	--	--	--	--
Good Impression	1.86	.86	1.35	.88	1.85	.71
Attention and Respect	1.86	.80	1.56	.96	2.02	.70
Confidence	1.90	.94	1.58	1.02	2.15	.66
Social Effectiveness	1.69	.94	--	--	--	--
RESPONSE						
Listening	2.11	.69	--	--	--	--
Flexibility	1.61	.96	1.50	1.03	--	--
Adaptability	1.88	.79	1.66	.85	2.20	.79
Reading Comprehension	--	--	1.63	.70	1.93	.76
PROBLEM ANALYSIS						
Organization	1.46	.79	--	--	1.85	.85
Concern for Details	--	--	1.41	.76	1.93	.74
Raising Questions	1.62	.81	--	--	--	--
Making Predictions	--	--	1.08	.76	--	--
DECISION-MAKING						
Considering Alternatives	1.60	1.00	1.22	.72	1.53	1.03
Decisiveness	1.62	.90	1.81	.86	2.06	.75
Value Clarification	--	--	--	--	1.62	.84

TABLE 3

Paired Comparisons (t-test) Per Criterion Across Exercises

	Group Discussion vs. In-Basket	In-Basket vs. Oral Presentation	Oral Presentation vs. Group Discussion
INITIATION			
Oral Expression	--	--	--
Clear Writing	--	--	--
Initiative	$t = 1.19$	--	--
INFLUENCE			
Persistence	--	--	--
Good Impression	$t = 3.2^{**}$	$t = -3.85^{**}$	$t = .177$
Attention and Respect	$t = 1.87$	$t = -2.87^{**}$	$t = -1.85$
Confidence	$t = 1.59$	$t = -3.42^{**}$	$t = -2.38^{*}$
Social Effectiveness	--	--	--
RESPONSE			
Listening	--	--	--
Flexibility	$t = 1.44$	--	--
Adaptability	$t = 1.34$	$t = -3.64^{**}$	$t = -2.19^{*}$
Reading Comprehension	--	$t = -2.34^{*}$	--
PROBLEM ANALYSIS			
Organization	--	--	$t = -2.32^{*}$
Concern for details	--	$t = -3.18^{**}$	--
Raising Questions	--	--	--
Making Predictions	--	--	--
DECISION-MAKING			
Considering Alternatives	$t = 1.91$	$t = -1.51$	$t = .23$
Decisiveness	$t = -1.26$	$t = -1.77$	$t = -2.81^{*}$
Value Clarification	--	--	--

* $p < .05$

** $p < .01$

TABLE 4

Correlation Coefficients Comparing Criteria
Across the Three Exercises

	Group Discussion vs. In-Basket	In-Basket vs. Oral Presentation	Oral Presentation vs. Group Discussion
INITIATION			
Oral Expression	--	--	--
Clear Writing	--	--	--
Initiative	.12	--	--
INFLUENCE			
Persistence	--	--	--
Good Impression	.22	.38**	.40**
Attention and Respect	.26	.27	.48**
Confidence	.12	.28	.47**
Social Effectiveness	--	--	--
RESPONSE			
Listening	--	--	--
Flexibility	.089	--	--
Adaptability	.009	.22	.12
Reading Comprehension	--	.17	--
PROBLEM ANALYSIS			
Organization	--	--	.096
Concern for details	--	-.067	--
Raising Questions	--	--	--
Making Predictions	--	--	--
DECISION-MAKING			
Considering Alternatives	-.067	-.13	.19
Decisiveness	.20	.27*	.21
Value Clarification	--	--	--

* $p < .05$

** $p < .01$

were not highly consistent across exercises, which suggests that certain skills may be a function of the specific situation in which they are demonstrated. One also wonders if the assessors do interpret the same capabilities differently in different situations, or even if the stimuli are equivalent in eliciting similar behaviors. That is, are the behaviors truly comparable? This question has implications for our understanding of students' ability to transfer competences to different situations.

In addition, we intercorrelated all criterion scores (n=36), disregarding exercise. The correlation matrix[1] indicated clusters of higher positive correlations between In-Basket criteria and Oral Presentation criteria. For example, all three criteria, Good Impression, Attention and Respect, and Confidence in the In-Basket correlated highly with each other, and formed a cluster with similar abilities in Oral Presentation. All criteria associated with the broader category of Decision-Making which include Considering Alternatives, Decisiveness and Value Clarification were highly intercorrelated in the In-Basket and formed a cluster with similar criteria in the Oral Presentation. These criteria under Decision-Making are also significantly correlated with Flexibility and Adaptability within the In-Basket exercise. These correlations suggest that situations where a person acts alone may elicit the same abilities. The In-Basket and Oral Presentation are individual exercises whereas the Group Discussion adds an interpersonal dimension to the same skill. Table 4, however, shows positive correlations between Oral Presentation and Group Discussion on behaviors which reflect personal style. Thus, personal style may be consistent irrespective of the exercise mode.

Still another issue of interest in the analysis of student performance on the Integrated Competence Seminar is the extent to which certain behavioral criteria represent behaviors that are easy vs. difficult for students to perform after two years in college. Which behaviors are performed by most students?

---

[1]This 36 X 36 intercorrelation matrix is available from the Office of Research and Evaluation.

Thus, the next step was to further examine the distribution of student performance on each criterion. As stated earlier, criteria are grouped under five broader categories: Initiation, Influence, Response, Problem Analysis and Decision-Making. Group Discussion has 13 criteria, the In-Basket has 12, and the Oral Presentation has 11. Criteria were not identical across each exercise, of course. Table 5 shows the percentage of students (n=48) who performed at each point on the scale per behavioral criterion for each exercise.

An examination of Table 5 shows that the students are rather evenly distributed across three of the points on the scale for most criteria, namely, As Often as Possible, Frequently, and At Least Once. In general small percentages of students responded at the Not at All point on the scale. These distributions confirm that the Integrated Competence Seminar may serve as an external criterion measure that will provide variability in performance that can be compared to that of other students, and to performance on other measures. Since relatively few percentages of students respond in the Not at All category, we may conclude that students do demonstrate the behaviors as opposed to not demonstrating them. That is, they have achieved the behaviors to some measureable degree.

There are, however, some behaviors that seem less likely to be demonstrated, and seem to stand out from all the rest. These are important because we may wish to review the stimulus to examine the extent to which the measure actually elicits or provides opportunity for demonstrating these behaviors. We might also examine the extent to which we would expect that these behaviors are likely to be exhibited after just two years in college. These criteria also may indicate some excellent discriminative behaviors that would add to the value of the technique for comparison purposes as an external criterion measure. We chose, somewhat arbitrarily, the figure of 10% in the "not at all" scale point at a cutoff for identifying criteria

TABLE 5

Percentage of Students (n=48) Who Performed at Each Scale Point Per Behavioral Criterion for Each Integrated Competence Seminar Exercise

	GROUP DISCUSSION				IN-BASKET				ORAL PRESENTATION			
	As often as possible	Frequently	At least once	Not at all	As often as possible	Frequently	At least once	Not at all	As often as possible	Frequently	At least once	Not at all
INITIATION												
Oral Expression	--	--	--	--	--	--	--	--	22.2	57.8	20.0	--
Clear Writing	--	--	--	--	22.9	39.6	22.9	14.6	--	--	--	--
Initiative	21.6[a]	37.8	37.8	2.7	12.5	43.8	27.1	16.7	--	--	--	--
INFLUENCE												
Persistence	24.3	43.2	24.3	8.1	--	--	--	--	--	--	--	--
Good Impression	10.8	64.9	18.9	5.4	6.3	43.8	29.2	20.8	18.8	47.9	33.3	--
Attention and Respect	13.5	59.5	27.0	--	18.8	33.3	33.3	14.6	25.5	51.1	23.4	--
Confidence	18.9	51.4	24.3	5.4	20.9	34.9	25.6	18.6	30.4	54.3	15.2	--
Social Effectiveness	27.8	33.3	27.8	11.1	--	--	--	--	--	--	--	--
RESPONSE												
Listening	29.7	59.5	10.8	--	--	--	--	--	--	--	--	--
Flexibility	21.6	35.1	40.5	2.7	20.8	27.1	33.3	18.8	--	--	--	--
Adaptability	29.7	45.9	21.6	2.7	12.5	54.2	20.8	12.5	39.6	45.8	10.4	4.2
Reading Comprehension	--	--	--	--	6.4	57.4	29.8	6.4	23.4	48.9	25.5	2.1
PROBLEM ANALYSIS												
Organization	8.1	48.6	37.8	5.4	--	--	--	--	22.9	45.8	25.0	6.3
Concern for Details	--	--	--	--	4.2	45.8	37.5	12.5	19.6	58.7	17.4	4.3
Raising Questions	13.5	37.8	37.8	10.8	--	--	--	--	--	--	--	--
Making Predictions	--	--	--	--	4.2	20.8	54.2	20.8	--	--	--	--
DECISION-MAKING												
Considering Alternatives	18.9	45.9	24.3	10.8	2.1	35.3	50.0	14.6	19.1	36.2	23.4	21.3
Decisiveness	16.2	40.5	37.8	5.4	22.9	41.7	29.2	6.3	29.2	50.0	18.8	2.1
Value Clarification	--	--	--	--	--	--	--	--	16.7	35.4	41.7	6.3

[a]Scores in percent.

that may be such discriminative criteria. That is, if 10% or more of the students responded at the Not At All scale point, we felt the behavior would be one that students found to be more difficult to demonstrate then others.

An examination of Table 5 indicates that there are three criteria in the Group Discussion that may be difficult: Social Effectiveness, Raising Questions and Considering Alternatives. There are ten such criteria in the In-Basket: Clear Writing, Initiative, Good Impression, Attention and Respect, Confidence, Flexibility, Adaptability, Concern for Details, Making Predictions, and Considering Alternatives. For the Oral Presentation, there is one such criterion: Considering Alternatives.

These results indicate that 10 of the 12 criteria on the In-Basket show at least 10% of the students responding at the Not At All scale point: the range is from 12.5% to 20.8%. This may indicate that the In-Basket exercise score does discriminate high and low performing students rather well.

The finding that not all students performed all of the criteria in the In-Basket is confirmed when we examine the number of items completed in the In-Basket exercise. The In-Basket contains seven items in all; the students are required to organize their time, set priorities, and attempt to complete all items. Table 6 shows the percentage of students completing a certain number of items.

TABLE 6

Percentage of Students (n=48) Completing Each of the Seven Items in the In-Basket Exercise

Number of Items Completed	Percentage
None of the seven	0.00
One of seven	87.5
Two of seven	95.8
Three of seven	83.3
Four of seven	72.9
Five of seven	62.5
Six of seven	50.0
Seven of seven	50.0

Since 50 percent of the students completed all seven items, and 50 percent did not, the number of items completed in the In-Basket exercise can be considered an additional discriminative measure. The fact that number of items completed shows variation, as do the behavioral criteria for the In-Basket, suggests that the number of items completed also may be considered as a measure of performance on the In-Basket.

If the number of items completed would be used, the instrument designers might examine the underlying meaning of this measure in relation to the behaviors the In-Basket is designed to elicit. For example, how does the number of items completed relate to criteria such as organization, setting priorities and difficulty of items?

In order to further examine the extent to which each of the behavioral criteria discriminate between high and low performing students, we correlated each student's score per criterion with each student's corresponding total exercise score (see Table 7). If all criteria correlate significantly positive with the corresponding exercise total score the criteria are assumed to be nondiscriminative.

An examination of the correlation coefficients in Table 7 shows that all criteria within each exercise correlate significantly positive with the corresponding exercise score. Thus, we may not conclude that the criteria mentioned earlier (at least 10% of the students did not demonstrate them) are good discriminative items.

An alternate interpretation of the correlation matrix suggests that assessors, while rating student performance, may not discriminate among the criteria. That is, assessors may form a judgment about a person's overall performance using the behavioral observations he or she records to create this holistic expert judgment. When the quantitative checklist is completed,

TABLE 7

Correlations Between Each Exercise Behavioral Criterion and Total Exercise Score (n=41)

	Group Discussion	In-Basket	Oral Presentation
INITIATION			
Oral Expression	--	--	.81
Clear Writing	--	.74	--
Initiative	.72	.79	--
INFLUENCE			
Persistence	.80	--	--
Good Impression	.80	.86	.84
Attention and Respect	.80	.82	.80
Confidence	.85	.79	.75
Social Effectiveness	.86	--	--
RESPONSE			
Listening	.75	--	--
Flexibility	.76	.90	--
Adaptability	.85	.86	.72
Reading Comprehension	--	.84	.81
PROBLEM ANALYSIS			
Organization	.79	--	.85
Concern for Details	--	.78	.76
Raising Questions	.66	--	--
Making Predictions	--	.70	--
DECISION-MAKING			
Considering Alternatives	.73	.65	.76
Decisiveness	.69	.76	.83
Value Clarification	--	--	.78

NOTE: All correlation coefficients are significant at the .001 level.

however, this expert judgment colors judgment on each of the individual criteria. This is known as the "halo effect." That is, if an assessor is likely to judge a student's overall performance as high, the assessor is also likely to judge the student's performance as "high" on each of the separate checklist criteria. This interpretation is not only supported by the significantly positive correlation coefficients between each exercise criterion and the corresponding exercise total score. We also obtained an intercorrelation matrix for each exercise that shows the correlation of each criterion with every other criterion. These intercorrelation matrices are depicted in Tables 8, 9, and 10. Examination of these matrices shows positive significant correlation coefficients for almost each comparison of a criterion with each of the others in the exercise. This pattern also seems to indicate that assessors are less likely to differentiate between students on a single criterion; rather, they do differentiate students on an overall exercise score, or on the ICS total score. One might speculate that assessors use the specific criteria to call forth the behavioral observations they made earlier, but that when they make a judgment per criterion, the judgment is affected by the overall concept of a student's performance (halo effect). Thus, students who receive an overall high rating are also rated high on all criteria.

Although we have some evidence for the operation of a halo effect in the criterion rating given by assessors, assessors do consistently agree with each other on an overall low performance or an overall high performance rating for each student. The correlation between the individual assessor rating and consensus for the Group Discussion exercise was .85 (see page 6). Further, assessors do distribute their ratings across the four scale points.

In sum, analysis of the specific criteria suggests that assessors make

TABLE 8

Intercorrelations Within Group Discussion

		GROUP1	GROUP2	GROUP3	GROUP4	GROUP5	GROUP6	GROUP7	GROUP8	GROUP9	GROUP10	GROUP11	GROUP12	GROUP13
Initiative	GROUP1	1.0000	.6201	.4618	.4947	.5964	.6269	.4908	.4332	.6634	.4692	.4564	.4228	.4474
		(42)	(42)	(42)	(42)	(42)	(41)	(42)	(42)	(42)	(42)	(42)	(42)	(42)
		S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .002	S= .001	S= .001	S= .001	S= .003	S= .001
Persistence	GROUP2	.6201	1.0000	.6032	.6139	.6489	.6906	.5438	.5846	.6331	.5368	.4374	.4770	.4546
		(42)	(43)	(43)	(43)	(43)	(42)	(43)	(43)	(43)	(43)	(43)	(43)	(43)
		S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .002	S= .001	S= .001
Good Impression	GROUP3	.4618	.6032	1.0000	.8992	.8890	.5670	.6619	.5206	.5333	.5821	.3642	.4030	.4534
		(42)	(43)	(43)	(43)	(43)	(42)	(43)	(43)	(43)	(43)	(43)	(43)	(43)
		S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .008	S= .004	S= .001
Attention and Respect	GROUP4	.4947	.6139	.8992	1.0000	.8896	.6094	.5816	.5878	.4966	.5121	.4263	.3431	.4527
		(42)	(43)	(43)	(43)	(43)	(42)	(43)	(43)	(43)	(43)	(43)	(43)	(43)
		S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .002	S= .012	S= .001
Confidence	GROUP5	.5964	.6489	.8890	.8896	1.0000	.6392	.5940	.5209	.6196	.6266	.3851	.4116	.6010
		(42)	(43)	(43)	(43)	(43)	(42)	(43)	(43)	(43)	(43)	(43)	(43)	(43)
		S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .005	S= .003	S= .001
Social Effectiveness	GROUP6	.6269	.6906	.5670	.6094	.6392	1.0000	.5658	.6942	.7720	.4855	.6864	.6363	.4921
		(41)	(42)	(42)	(42)	(42)	(42)	(42)	(42)	(42)	(42)	(42)	(42)	(42)
		S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001
Listening	GROUP7	.4908	.5438	.6619	.5816	.5940	.5658	1.0000	.5947	.5417	.6285	.5374	.3736	.3740
		(42)	(43)	(43)	(43)	(43)	(42)	(43)	(43)	(43)	(43)	(43)	(43)	(43)
		S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .007	S= .007
Flexibility	GROUP8	.4332	.5846	.5206	.5878	.5209	.6942	.5947	1.0000	.6089	.4373	.5433	.4807	.3314
		(42)	(43)	(43)	(43)	(43)	(42)	(43)	(43)	(43)	(43)	(43)	(43)	(43)
		S= .002	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .002	S= .001	S= .001	S= .015
Adaptability	GROUP9	.6634	.6331	.5333	.4966	.6196	.7720	.5417	.6089	1.0000	.6524	.5562	.6590	.4715
		(42)	(43)	(43)	(43)	(43)	(42)	(43)	(43)	(43)	(43)	(43)	(43)	(43)
		S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001
Organization	GROUP10	.4692	.5368	.5821	.5121	.6266	.4855	.6285	.4373	.6524	1.0000	.4547	.6818	.6780
		(42)	(43)	(43)	(43)	(43)	(42)	(43)	(43)	(43)	(43)	(43)	(43)	(43)
		S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .002	S= .001	S= .001	S= .001	S= .001	S= .001
Raising Questions	GROUP11	.4564	.4374	.3642	.4263	.3851	.6864	.5374	.5433	.5562	.4547	1.0000	.4842	.2604
		(42)	(43)	(43)	(43)	(43)	(42)	(43)	(43)	(43)	(43)	(43)	(43)	(43)
		S= .001	S= .002	S= .008	S= .002	S= .015	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .046
Considering Alternatives	GROUP12	.4228	.4770	.4030	.3431	.4116	.6363	.3736	.4807	.6590	.6818	.4842	1.0000	.6504
		(42)	(43)	(43)	(43)	(43)	(42)	(43)	(43)	(43)	(43)	(43)	(43)	(43)
		S= .003	S= .001	S= .004	S= .012	S= .003	S= .001	S= .007	S= .001	S= .001	S= .001	S= .001	S= .001	S= .001
Decisiveness	GROUP13	.4474	.4546	.4534	.4527	.6010	.4921	.3740	.3314	.4715	.6780	.2604	.6504	1.0000
		(42)	(43)	(43)	(43)	(43)	(42)	(43)	(43)	(43)	(43)	(43)	(43)	(43)
		S= .001	S= .001	S= .001	S= .001	S= .001	S= .001	S= .007	S= .015	S= .001	S= .001	S= .046	S= .001	S= .001

TABLE 9

Intercorrelations Within In-Basket

	BASKET1	BASKET2	BASKET3	BASKET4	BASKET5	BASKET6	BASKET7	BASKET8	BASKET9	BASKET10	BASKET11	BASKET12
Clear Writing	1.0000 (48) S= .001	.5670 (48) S= .001	.4844 (48) S= .001	.5323 (48) S= .001	.7129 (43) S= .001	.5009 (48) S= .001	.6600 (48) S= .001	.6160 (47) S= .001	.4162 (48) S= .002	.3133 (48) S= .015	.2150 (48) S= .071	.6302 (48) S= .001
Initiative	.5670 (48) S= .001	1.0000 (48) S= .001	.6020 (48) S= .001	.5003 (48) S= .001	.6307 (43) S= .001	.6374 (48) S= .001	.5730 (48) S= .001	.6199 (47) S= .001	.5588 (48) S= .001	.5087 (48) S= .001	.3283 (48) S= .011	.5506 (48) S= .001
Good Impression	.4844 (48) S= .001	.6020 (48) S= .001	1.0000 (48) S= .001	.6320 (48) S= .001	.5577 (43) S= .001	.8023 (48) S= .001	.6331 (48) S= .001	.6847 (47) S= .001	.6539 (48) S= .001	.6435 (48) S= .001	.5353 (48) S= .001	.4480 (48) S= .001
Attention and Respect	.5323 (48) S= .001	.5003 (48) S= .001	.6320 (48) S= .001	1.0000 (48) S= .001	.7307 (43) S= .001	.6303 (48) S= .001	.5645 (48) S= .001	.6493 (47) S= .001	.5673 (48) S= .001	.4811 (48) S= .001	.4523 (48) S= .001	.6626 (48) S= .001
Confidence	.7129 (43) S= .001	.6307 (43) S= .001	.5577 (43) S= .001	.7307 (43) S= .001	1.0000 (43) S= .001	.5428 (43) S= .001	.6578 (43) S= .001	.6299 (42) S= .001	.5200 (43) S= .001	.4347 (43) S= .002	.3695 (43) S= .007	.6184 (43) S= .001
Flexibility	.5009 (48) S= .001	.6374 (48) S= .001	.8023 (48) S= .001	.6303 (48) S= .001	.5428 (43) S= .001	1.0000 (48) S= .001	.6725 (48) S= .001	.7024 (47) S= .001	.7259 (48) S= .001	.6184 (48) S= .001	.5288 (48) S= .001	.5355 (48) S= .001
Adaptability	.6600 (48) S= .001	.5730 (48) S= .001	.6331 (48) S= .001	.5645 (48) S= .001	.6578 (43) S= .001	.6725 (48) S= .001	1.0000 (48) S= .001	.6837 (47) S= .001	.5382 (48) S= .001	.4628 (48) S= .001	.5035 (48) S= .001	.6574 (48) S= .001
Reading Comprehension	.6160 (47) S= .001	.6199 (47) S= .001	.6847 (47) S= .001	.6493 (47) S= .001	.6299 (42) S= .001	.7024 (47) S= .001	.6837 (47) S= .001	1.0000 (47) S= .001	.7553 (47) S= .001	.4159 (47) S= .002	.3689 (47) S= .005	.6248 (47) S= .001
Concern for Details	.4162 (48) S= .002	.5588 (48) S= .001	.6539 (48) S= .001	.5673 (48) S= .001	.5200 (43) S= .001	.7259 (48) S= .001	.5382 (48) S= .001	.7553 (47) S= .001	1.0000 (48) S= .001	.5542 (48) S= .001	.4771 (48) S= .001	.4719 (48) S= .001
Making Predictions	.3133 (48) S= .015	.5087 (48) S= .001	.6435 (48) S= .001	.4811 (48) S= .001	.4347 (43) S= .002	.6184 (48) S= .001	.4628 (48) S= .001	.4159 (47) S= .002	.5542 (48) S= .001	1.0000 (48) S= .001	.7333 (48) S= .001	.3759 (48) S= .004
Considering Alternatives	.2150 (48) S= .071	.3283 (48) S= .011	.5353 (48) S= .001	.4523 (48) S= .001	.3695 (43) S= .007	.5288 (48) S= .001	.5035 (48) S= .001	.3689 (47) S= .005	.4771 (48) S= .001	.7333 (48) S= .001	1.0000 (48) S= .001	.4103 (48) S= .002
Decisiveness	.6302 (48) S= .001	.5506 (48) S= .001	.4480 (48) S= .001	.6626 (48) S= .001	.6184 (43) S= .001	.5355 (48) S= .001	.6574 (48) S= .001	.6248 (47) S= .001	.4719 (48) S= .001	.3759 (48) S= .004	.4103 (48) S= .002	1.0000 (48) S= .001

TABLE 10

Intercorrelations Within Oral Presentation

	ORAL1	ORAL2	ORAL3	ORAL4	ORAL5	ORAL6	ORAL7	ORAL8	ORAL9	ORAL10	ORAL11
Oral Expression	1.0000 (45) S= .001	.7584 (45) S= .001	.7507 (44) S= .001	.7250 (43) S= .001	.5083 (44) S= .001	.5503 (45) S= .001	.5687 (45) S= .001	.5804 (43) S= .001	.4755 (44) S= .001	.5735 (45) S= .001	.5393 (45) S= .001
Good Impression	.7584 (45) S= .001	1.0000 (48) S= .001	.7845 (47) S= .001	.6467 (46) S= .001	.6137 (47) S= .001	.5772 (48) S= .001	.6299 (48) S= .001	.5378 (46) S= .001	.5059 (47) S= .001	.6877 (48) S= .001	.5091 (48) S= .001
Attention and Respect	.7507 (44) S= .001	.7845 (47) S= .001	1.0000 (47) S= .001	.6937 (46) S= .001	.5755 (46) S= .001	.6026 (47) S= .001	.5474 (47) S= .001	.4709 (46) S= .001	.4873 (47) S= .001	.6018 (47) S= .001	.4135 (47) S= .002
Confidence	.7250 (43) S= .001	.6467 (46) S= .001	.6937 (46) S= .001	1.0000 (46) S= .001	.5160 (45) S= .001	.4817 (46) S= .001	.5575 (46) S= .001	.3974 (45) S= .003	.4250 (46) S= .002	.5426 (46) S= .001	.3754 (46) S= .005
Reading Comprehension	.5893 (44) S= .001	.6137 (47) S= .001	.5755 (46) S= .001	.5160 (45) S= .001	1.0000 (47) S= .001	.5526 (47) S= .001	.7475 (47) S= .001	.7177 (45) S= .001	.5986 (46) S= .001	.5667 (47) S= .001	.5448 (47) S= .001
Adaptability	.5503 (45) S= .001	.5772 (48) S= .001	.6026 (47) S= .001	.4817 (46) S= .001	.5526 (47) S= .001	1.0000 (48) S= .001	.5789 (48) S= .001	.4674 (46) S= .001	.3810 (47) S= .004	.4722 (48) S= .001	.5311 (48) S= .001
Organization	.5687 (45) S= .001	.6299 (48) S= .001	.5474 (47) S= .001	.5575 (46) S= .001	.7475 (47) S= .001	.5789 (48) S= .001	1.0000 (48) S= .001	.7876 (46) S= .001	.5460 (47) S= .001	.6440 (48) S= .001	.6062 (48) S= .001
Concern for Details	.5804 (43) S= .001	.5378 (46) S= .001	.4709 (46) S= .001	.3974 (45) S= .003	.7177 (45) S= .001	.4674 (46) S= .001	.7876 (46) S= .001	1.0000 (46) S= .001	.4516 (46) S= .001	.5179 (46) S= .001	.4876 (46) S= .001
Considering Alternatives	.4755 (44) S= .001	.5059 (47) S= .001	.4873 (47) S= .001	.4250 (46) S= .002	.5986 (46) S= .001	.3810 (47) S= .004	.5460 (47) S= .001	.4516 (46) S= .001	1.0000 (47) S= .001	.6686 (47) S= .001	.6197 (47) S= .001
Decisiveness	.5735 (45) S= .001	.6877 (48) S= .001	.6018 (47) S= .001	.5426 (46) S= .001	.5667 (47) S= .001	.4722 (48) S= .001	.6440 (48) S= .001	.5179 (46) S= .001	.6686 (47) S= .001	1.0000 (48) S= .001	.6405 (48) S= .001
Value Clarification	.5393 (45) S= .001	.5091 (48) S= .001	.4135 (47) S= .002	.3754 (46) S= .005	.5448 (47) S= .001	.5311 (48) S= .001	.6062 (48) S= .001	.4876 (46) S= .001	.6197 (47) S= .001	.6405 (48) S= .001	1.0000 (48) S= .001

reliable judgments about the overall exercise score on the Integrated Competence Seminar, and recognize, evaluate and record the individual differences between students on this overall rating. It may be that assessors use the specific criteria as guides to "what to look for," to record behavioral observations, and to detail the various aspects of the student's performance that comprise the general overall concept the assessor forms about the student's overall performance on the Integrated Competence Seminar. Just as the student is expected to integrate her performance on this simulated assessment technique, so the assessor integrates the information from the pattern of behavioral observations he or she makes, to form a highly reliable judgment that discriminates students' individual differences in performance.

The forty-eight students assessed also demonstrated that, as a group, almost all performed adequately on the instrument.

Validating the Integrated Competence Seminar in Relation to Other Student Performance Measures in the Longitudinal Study

As mentioned earlier, the forty-eight students comprising the sample used for the preceding analyses were part of a longitudinal sample of students participating in a study of college outcomes. The Integrated Competence Seminar was employed as a generic measure of student performance in Alverno's efforts to validate its outcome-centered curriculum (Mentkowski & Doherty, 1977, 1983; Mentkowski, 1980). Several strategies for validating outcomes were developed into a comprehensive validation model. The model incorporates various research and evaluation methods, with the ultimate goals

of establishing program validity, contributing to program development, and developing a picture of adult learning and development that can be used to improve instruction and assessment in liberal education settings. Figure 5 displays the several components of the validation model. The results reported below link performance on the Integrated Competence Seminar assessment to other internal college performance variables (e.g., competence level units and credit hours), to performance characteristics, and to the external criterion measures of human potential. Other papers present results linking the human potential measures to the performance characteristics rating (Assessment Committee/Office of Research and Evaluation, 1983) and to the internal college performance measures (Mentkowski & Strait, 1983).

Longitudinal data were collected on two consecutive classes entering in 1976 and 1977. While the ICS was administered on a single occasion, at the midpoint of the student's education, the measures of human potential, identified in Table 11, were administered on three occasions: at entrance, two years later, and three and one-half years after entrance. The second longitudinal assessment roughly coincided with the midpoint administration of the ICS so that the relationships between ICS performance scores and measures of human potential can be examined in the three temporal relationships illustrated in Figure 6. Figure 6 also indicates when other college performance variables were collected. Competence level units, credits and number of semesters were tabulated at the time of the second and third longitudinal assessments; the Six Performance Characteristics rating for one student cohort occurred twice, at the end of the third and fourth academic years of the study; for the other cohort three times, at the end of the second, third and fourth years.

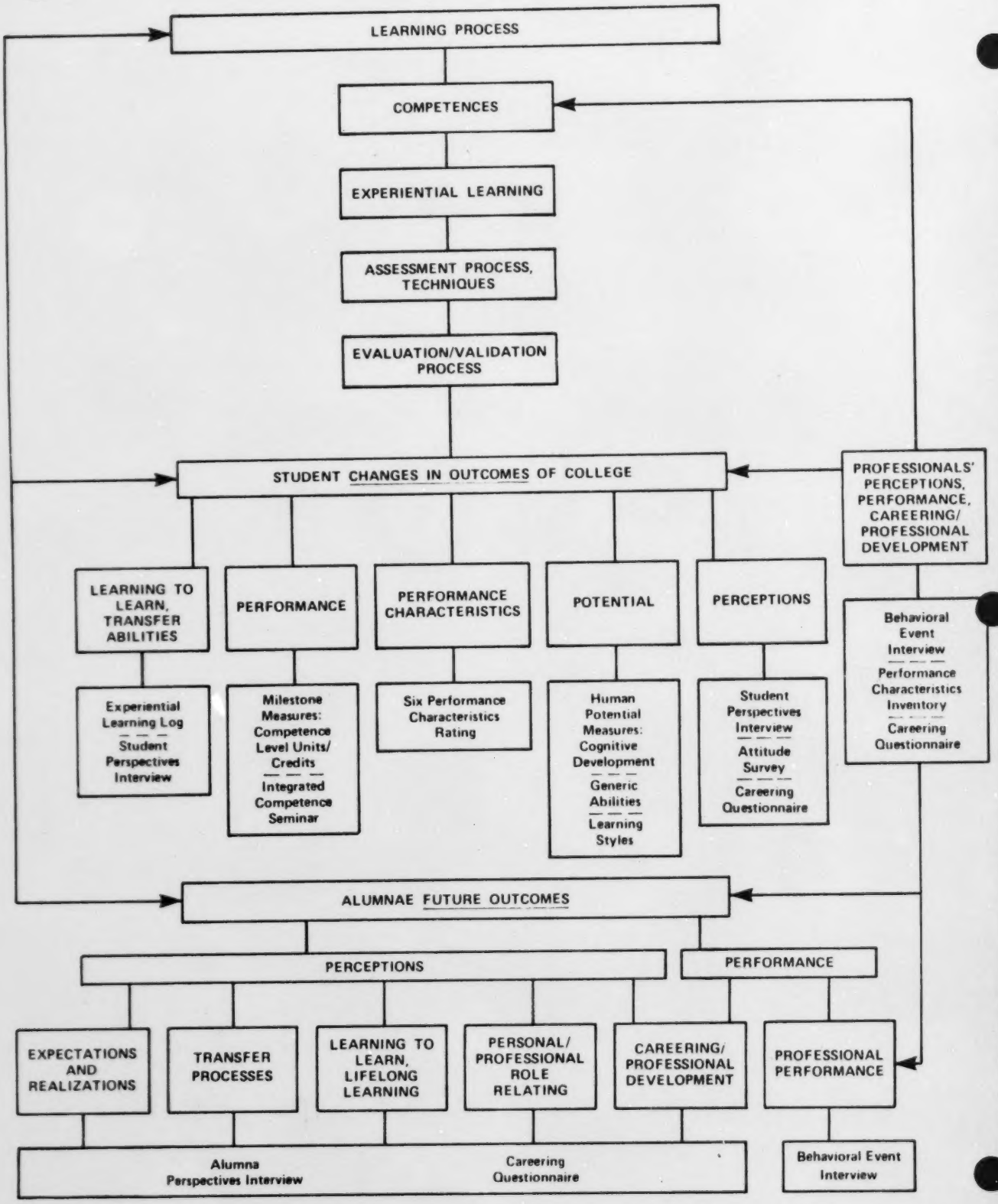


Figure 5. Components of a validation model for the Alverno learning process with external validation instruments.

Table 11.

Measures of Human Potential From the Longitudinal Study [1]

HUMAN POTENTIAL MEASURES: COGNITIVE DEVELOPMENT

Test of Cognitive Development (Renner et al., 1976; after Piaget)

By having a student work a series of problems and provide reasons for his or her answers, this instrument measures a student's cognitive activity based on Piaget's stages of cognitive development.

Sentence Completion Test (Loevinger, et al., 1970)

This instrument provides a measure of an individual's stage of ego development. "Ego" here is defined as one's "style of life," the unity of personality, individuality, the method of facing problems, opinion about one's self and the problems of life, and the whole attitude toward making choices in all life spheres.

Defining Issues Test (Rest, 1979)

Rest's instrument (based on Kohlberg's theory of moral development) provides a measure of an individual's moral development in a recognition task by analyzing the relative importance attributed by a person to principled moral considerations. A person attributes importance to several reasons given for resolving a particular moral dilemma, and then rank orders them.

Measure of Vocational, Educational, and Personal Issues (Knefelkamp, 1974; Widick, 1975; now titled: Measure of Intellectual Development; after Perry)

This measure of the Perry scheme of intellectual and ethical development asks students to write three essays on their best class, a major decision and their career. It assesses the progress the college student makes toward movement on the Perry scheme.

HUMAN POTENTIAL MEASURES: GENERIC ABILITIES

Test of Thematic Analysis* (Winter, 1976)

This instrument consists of two sets of stories students are asked to compare thematically. This "thematic analysis" is scored according to twelve categories of critical thinking. This test is based on an understanding of cognitive development defined as the ability to analyze new information and to synthesize new concepts based on this information, and reflects the ability to integrate information into one's own cognitive structure. As the cognitive structure grows, so does the ability to think critically, to make a cogent argument and to reason inductively.

*Available from McBer and Company.

Table 11. continued

Picture Story Exercise* (Scored for Stages of Adaptation (Stewart, 1977), Self-Definition (Stewart & Winter, 1974); and Achievement (McClelland, et al., 1953, Affiliation (Atkinson, 1958), and Power (Winter, 1973) motives.)

This instrument, modeled on the Thematic Apperception Test, may be used to assess a variety of abilities. The instrument requires the student to write narratives to six pictures. One is "self-definition" which encompasses the way one thinks about the world and one's self, the way one reacts to new information, and the way one behaves (Stewart and Winter, 1974). People with high cognitive initiative are not only able to think clearly, but also to reason from problem to solution, and to propose and take effective action on their own. This instrument is also used to assess Need for Achievement (McClelland, et al., 1953), Affiliation (Atkinson, 1958), Power (Winter, 1973), and Stages of Adaptation, a measure of ego development (created by Stewart, 1977).

Watson-Glaser Critical Thinking Appraisal (Watson and Glaser, 1964)

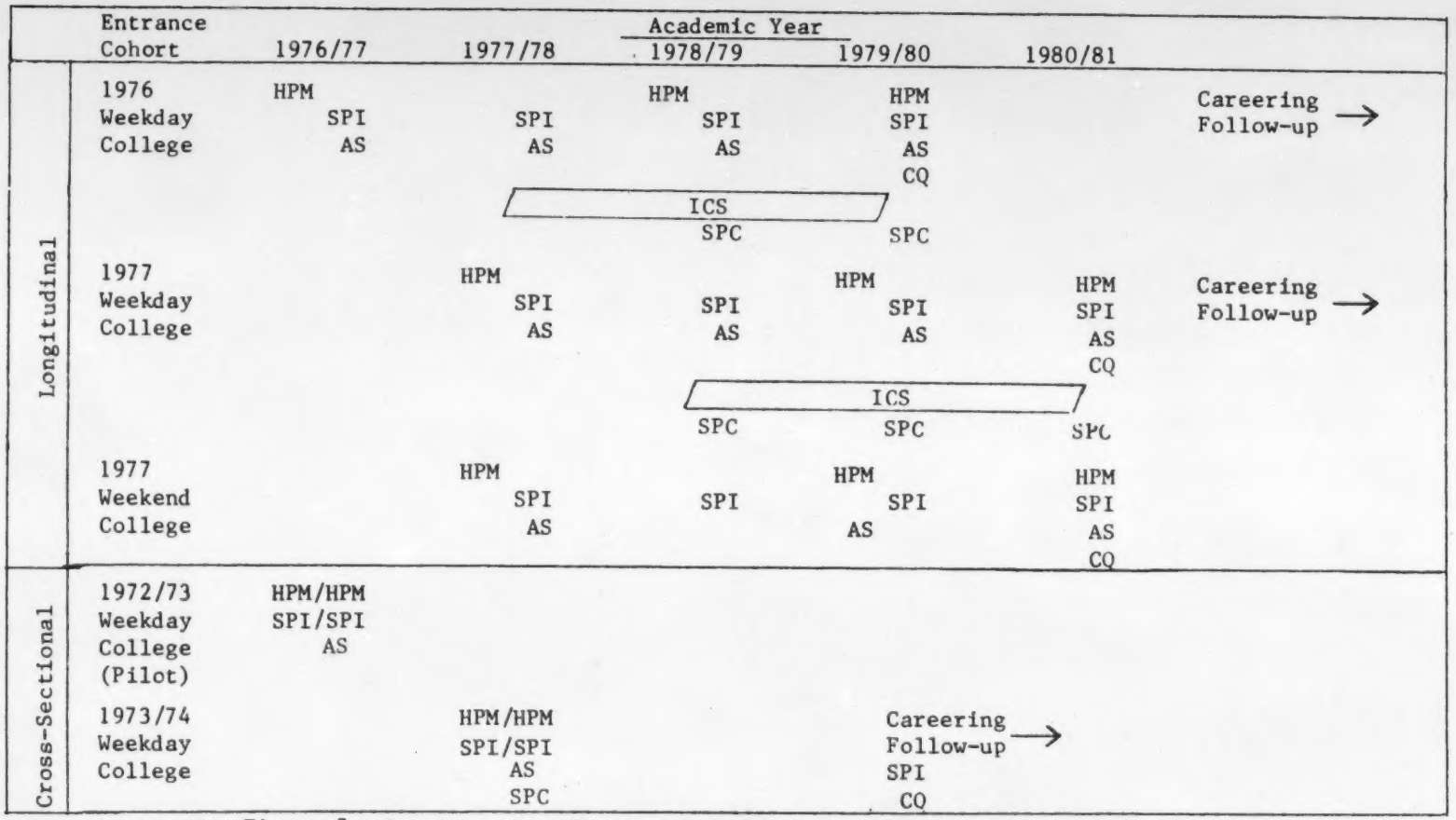
This instrument measures several components of critical thinking: Inference, Recognition of Assumptions, Deduction.

Learning Style Inventory* (Kolb, 1976)

The Learning Style Inventory is a measure of individual learning styles which affect decision-making and problem-solving. The four styles are Concrete Experience, Reflective Observation, Abstract Conceptualization, and Active Experimentation. The instrument requires the student to rank order descriptive statements about her mode of learning.

*Available from McBer and Company.

[1]For a more detailed description of the measures and their use in the longitudinal study, see Mentkowski, M. & Strait, M. A Longitudinal Study of Student Change in Cognitive Development and Generic Abilities in an Outcome-Centered Liberal Arts Curriculum. Milwaukee, Wi: Alverno Productions, 1983.



Note. See Figure 5 for overview of components of the program validation model with measures. Student Perspectives Interview (SPI) data were collected on a subsample of students participating in the administration of the Human Potential Measures (HPM), but all completed the Attitude Survey (AS) and Careering Questionnaire (CQ). All Weekday College students completed the Integrated Competence Seminar (ICS) and were rated by faculty on the Six Performance Characteristics (SPC).

Figure 6. Design for the administration of human potential measures and student perception measures for longitudinal and cross-sectional studies of student outcomes.

The major difference between the data analyzed and presented in this report concerning the human potential measure outcomes, and the data presented in the detailed study of student change (Mentkowski & Strait, 1983), is the absence, in the present study, of the Weekend College Cohort. The alternative Weekend College timeframe attracts older, working women whereas the Weekday College timeframe attracts the traditional age, less experienced college student. Consequently, the student population represented in this report is much more homogeneous with respect to age and other background and program characteristics.

The two consecutive entering classes comprising the present sample may be described in similar terms. At entrance, approximately 90% of the sample were single, about 80% were 17 to 19 years old, and about 70% identified themselves as Catholic. In terms of measured program variables, about two-thirds were commuting, and virtually all were attending college fulltime. Two-thirds of the students identified themselves at entrance as majoring in nursing, with the remainder scattered through Fine Arts, Education, and traditional liberal arts majors.

Analysis of ICS Scores for the Weekday College Longitudinal Sample

The combined classes provided a longitudinal sample of 136 students. Integrated Competence Seminar scores were available for 131 students. The distributions of students' scores per exercise and the ICS total score are presented in Figures 7, 8, 9, and 10. As on the previously described case (Figures 1, 2, 3, and 4), these graphs show that the highest scores were obtained on the Group Discussion exercise, and the lowest scores were obtained on the In-Basket exercise.

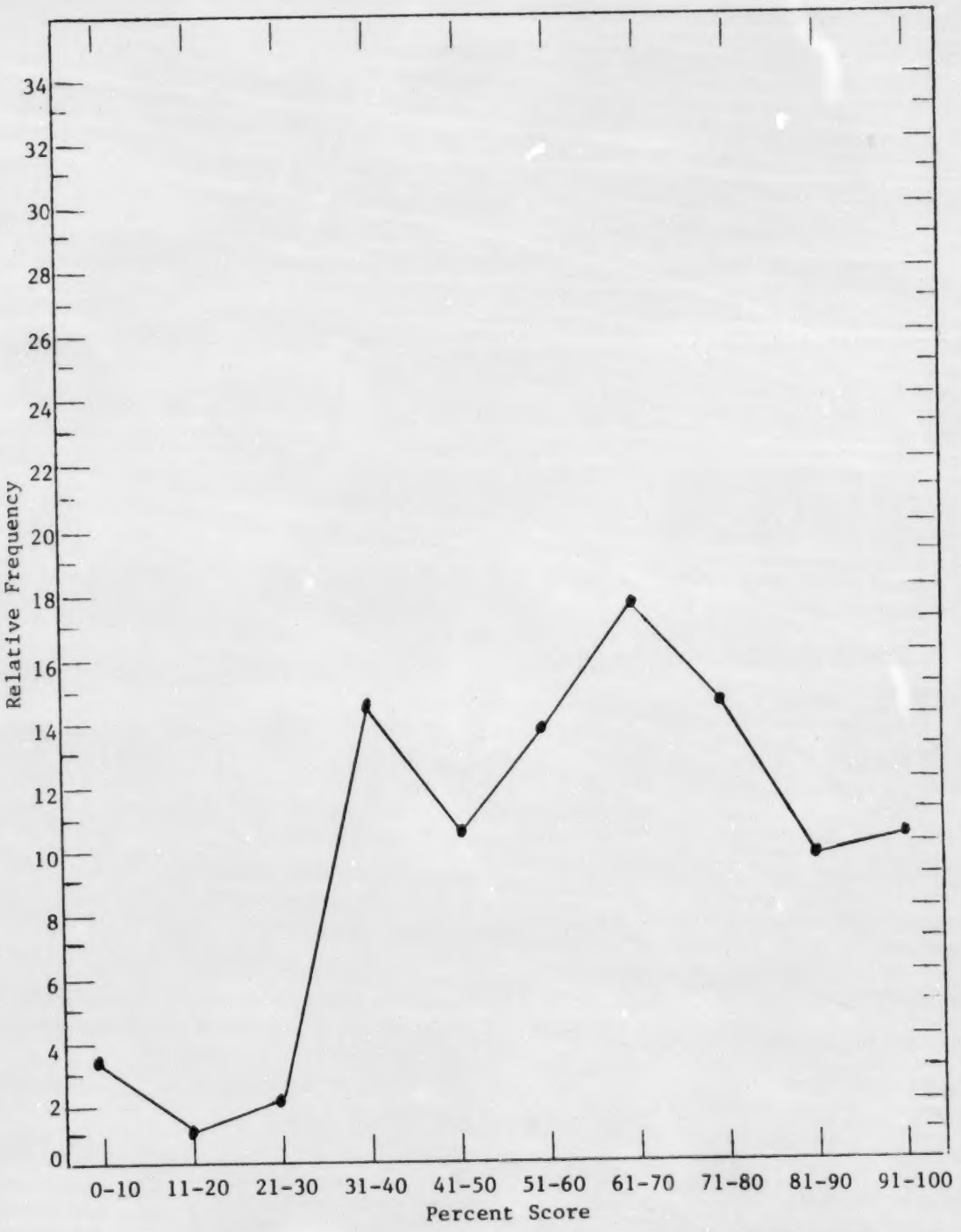


Figure 7. Relative Frequency Distribution of Students' Scores on the Group Discussion Exercise for the Longitudinal Study

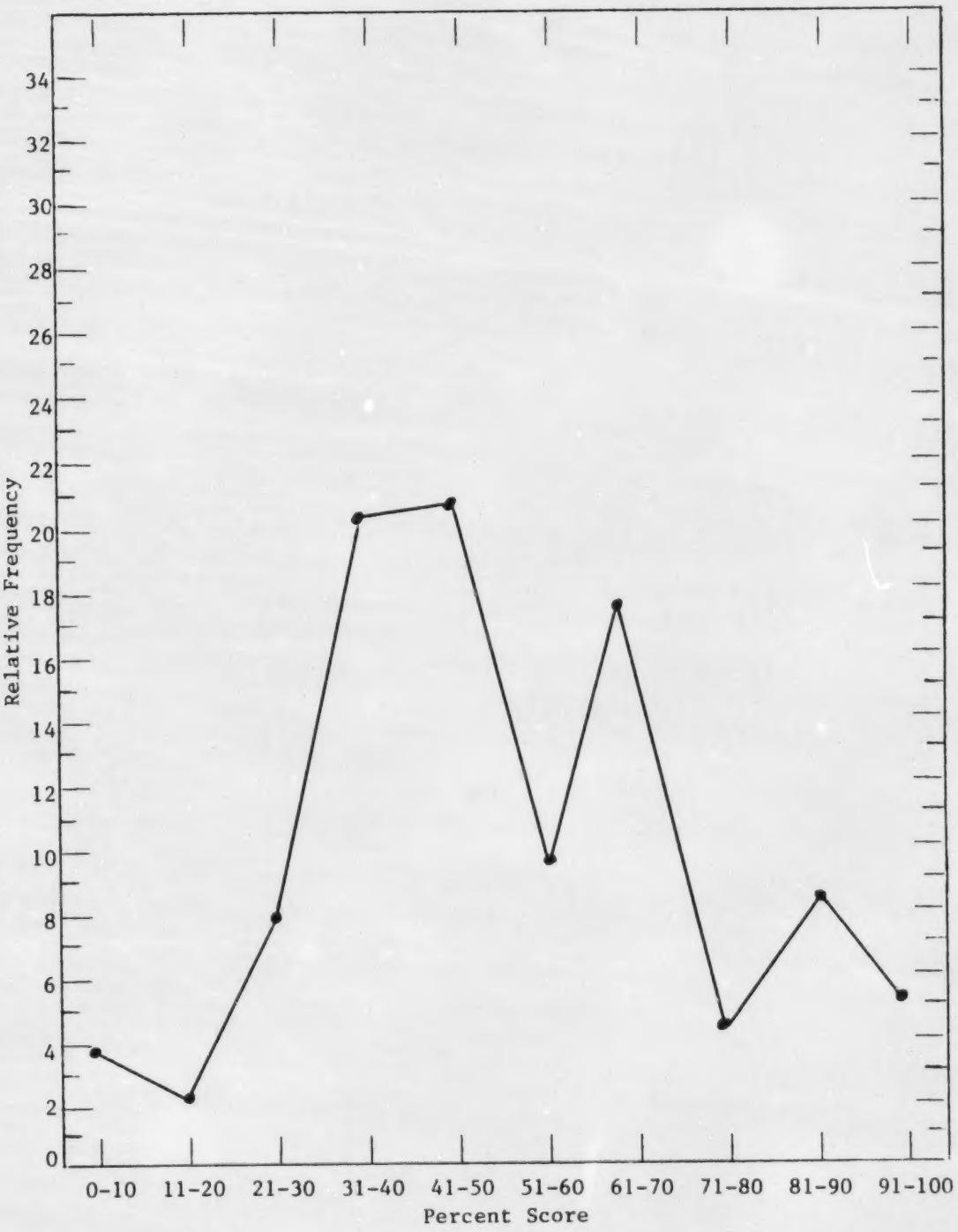


Figure 8. Relative Frequency Distribution of Students' Scores on the In-Basket Exercise for the Longitudinal Study

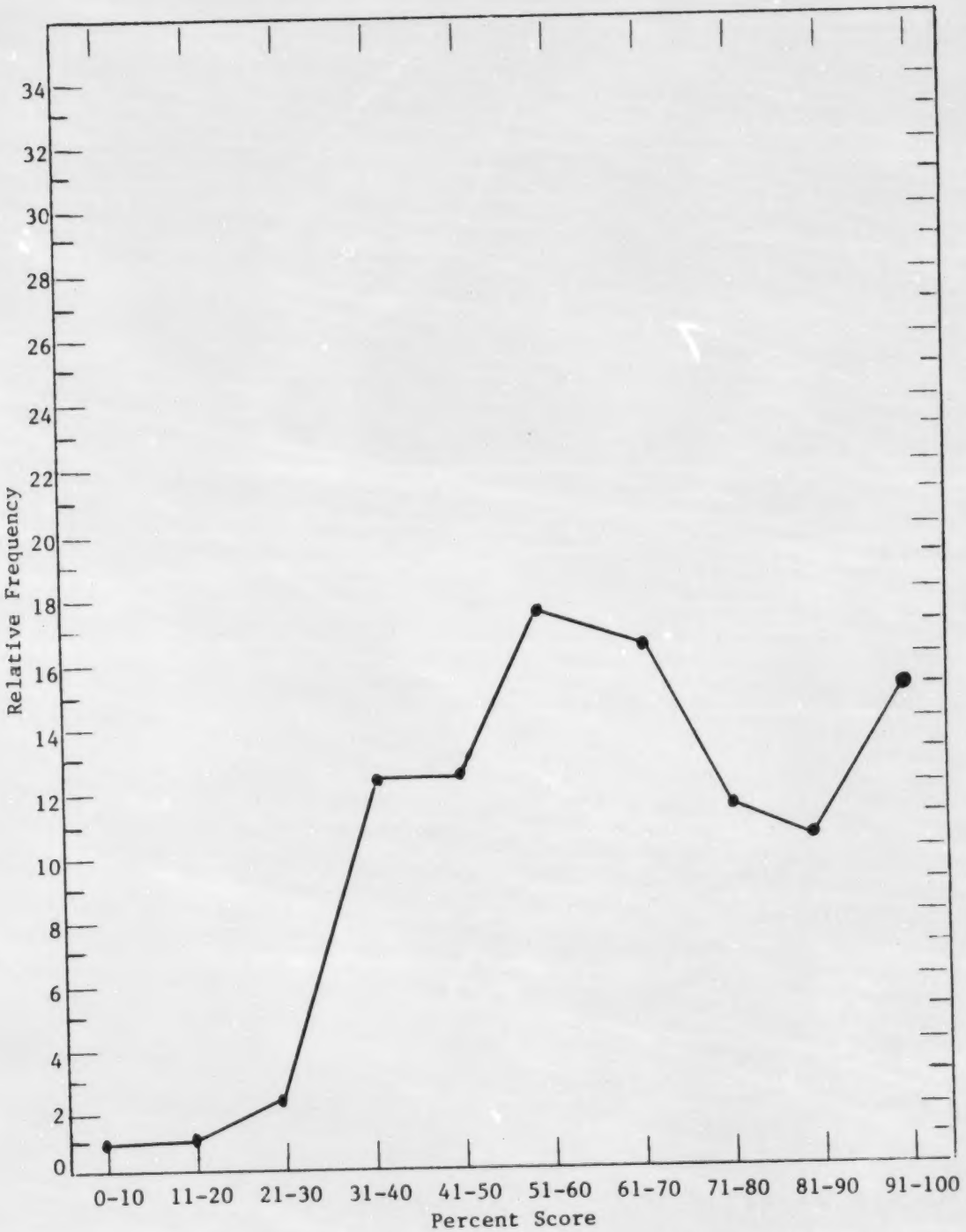


Figure 9. Relative Frequency Distribution of Students' Scores on the Oral Presentation Exercise for the Longitudinal Study

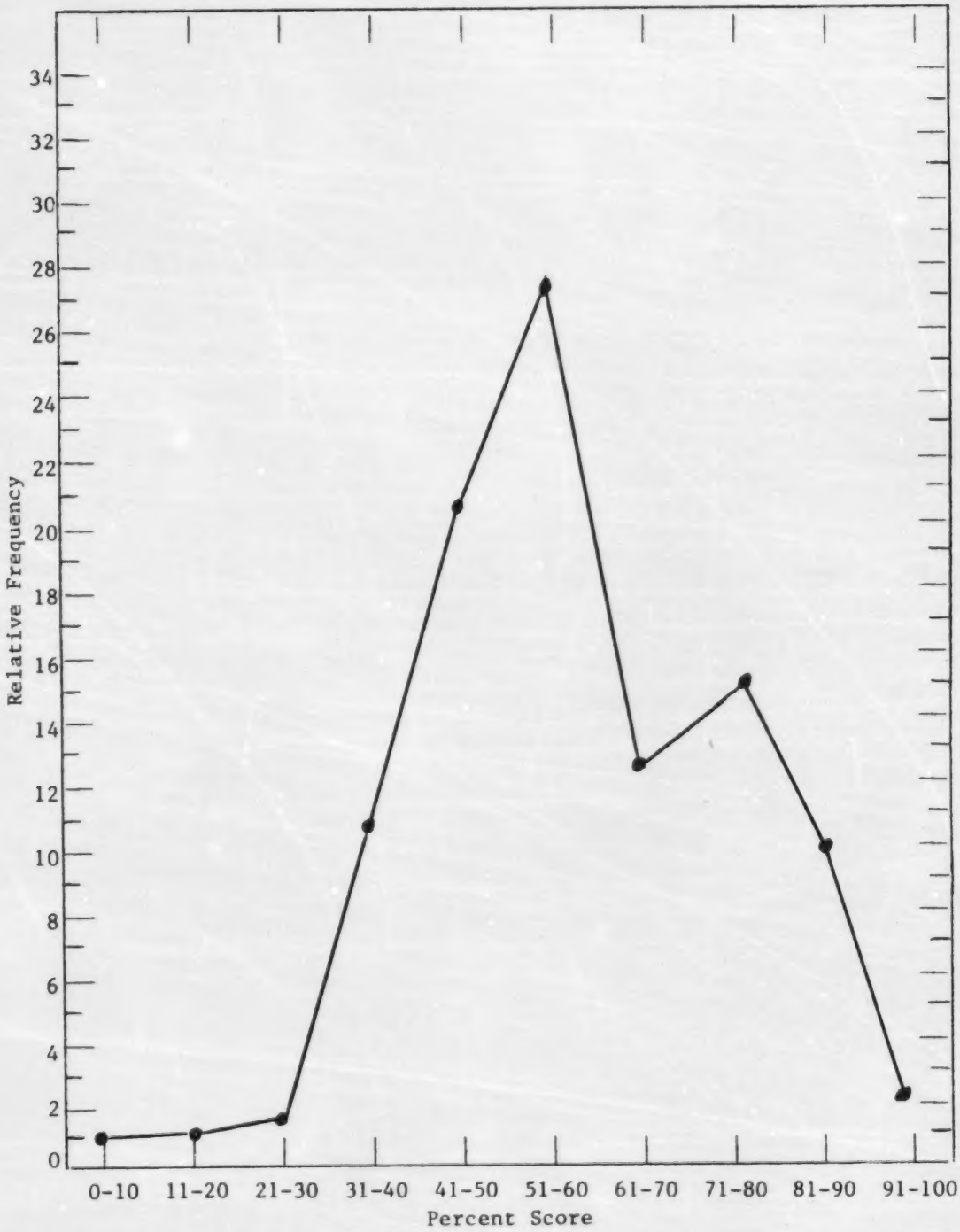


Figure 10. Relative Frequency Distribution of Students' Scores on the Integrated Competence Seminar for the Longitudinal Study

In the larger sample, there were more high scores on each exercise, but the total score distribution discloses that there were few students who received the highest scores on all three exercises. Correlation coefficients between the various exercises were all statistically significant. For In-Basket and Oral Presentation $r = .399$, $p < .001$; for In-Basket and Group Discussion $r = .260$, $p < .001$; for Oral Presentation and Group Discussion $r = .444$, $p < 001$. This would support that the exercises elicit different skills, assuming reliable measures, but that they share similar aspects as well.

Relationships Between ICS and
Student and Program Characteristics

There were few statistically significant relationships between ICS performance scores and measured background and program variables. Specifically, there were no significant correlations between ICS performance and age, parents' education or occupation, or high school grades. There were small but significant correlations between some ICS exercise scores or total score and variables recording religious affiliation and marital status. Students who were not identified as Catholic or who were not single tended to receive higher performance scores. Scores on the In-Basket exercise and the Total Score were significantly correlated with religious affiliation ($r = .203$, $n = 129$, $p = .011$, and $r = .179$, $n = 129$, $p = .021$, respectively). Marital status was correlated with performance scores on the Oral Presentation exercise ($r = .160$, $n = 131$, $p = .034$) and with Total Score ($r = .172$, $n = 131$, $p = .025$).

Considering the highly disproportional representations of subgroups and the high intercorrelations among background variables, it would be inappropriate to claim any theoretical association between religious affiliation

or marital status and performance on the ICS. The correlations suggest rather a demographic factor based on the local context of Alverno College. In the Weekday College, the typical student was of traditional age, single, Catholic, and a fulltime student. The observed correlations between background variables and performance on the ICS indicate that a student who was atypical with respect to the local context was likely to receive higher scores.

Among program variables, fulltime status was correlated with Oral Presentation and Total Score, but this relationship was redundant with relationships observed between ICS scores and background variables. There was a small but significant relationship between Oral Presentation scores and selected major field ($R = .314$, $F\ (6,123) = 2.24$, $p = .044$). Students majoring in behavioral sciences had the highest mean score, and students majoring in art the lowest. The disproportionately large group of nursing majors averaged below behavioral science, natural science, and education majors, but above art, arts and humanities, and undecided majors. There were no significant correlations between ICS performance scores and student residence.

The two entering class cohorts were not significantly different from one another on ICS performance scores; in fact, the averages were nearly identical. As partial confirmation of the typical/atypical dichotomy suggested above, students from the pooled entrance cohorts were reclassified as typical if they were single, Catholic, and fulltime; atypical if they were not all three. Sixty-four percent met all three criteria and were classified as typical. Based on this categorization, there were significant differences in performance on the ICS In-Basket exercise and the Total Score ($t\ (129) =$

2.47, p = .015, and t (129) = 2.55, p = .012, respectively), with atypical students scoring higher. Differences between typical and atypical students were in the same direction for Group Discussion and Oral Presentation exercise scores, but not statistically significant.

More important for present purposes than what group scored higher is that fact that ICS scores did correlate with background and program variables. Only the Group Discussion exercise score did not correlate with any background or program variable.

Relationships Between ICS and Other College Performance Measures

Table 12 presents the partial correlations between performance on the three ICS exercises, and the ICS Total Score, and the other college performance measures, controlling for variability due to background and program variables. Due to the variability in administration of the Integrated Competence Seminar to students in the two entering classes, it is not possible to place all indices in a clear temporal order. Generally, the ICS scores may be regarded as concurrent with credit, semester, and competence level unit information collected to coincide with the second longitudinal assessment of human potential measures. The first Six Performance Characteristics rating would also have occurred near the same time as ICS administrations. The data collected to coincide with the third and final longitudinal assessment would clearly have been collected after the ICS was administered.

In the combined Weekday College samples, the number of semesters attended by the second longitudinal assessment was correlated with major

Table 12.

Partial Correlations of ICS Scores With Other College Performance Measures, Controlling For Background and Program Covariates

	Exercises			Total Score (2,3,7)
	Group Discussion	In-Basket (2)	Oral Presentation (3,7,8)	
Concurrent Measures				
Number of Semesters (8)	.042 (127)	-.252** (124)	-.014 (126)	-.114 (123)
Credit Hours (1,7,8)	-.187* (127)	.013 (122)	.047 (125)	-.070 (122)
Competence Level Units (4,5,7,8)	-.011 (92)	-.003 (91)	-.082 (93)	-.044 (91)
Six Performance Characteristics Rating (4,6)	.134 (98)	.180* (97)	.082 (96)	.139 (95)
Later Measures				
Number of Semesters (-)	.296*** (127)	.162* (126)	.082 (126)	.223** (124)
Credit Hours (9)	.103 (127)	.120 (125)	.030 (125)	.117 (123)
Competence Level Units (10)	.044 (126)	.117 (124)	-.053 (124)	.057 (122)
Six Performance Characteristics Rating (4,11)	.042 (98)	-.005 (97)	.073 (96)	.044 (95)

*p < .05
**p < .01
***p < .001

Note. Numbers in parentheses in cells are sample sizes. Numbers in parentheses in headings are covariates as follows:

1 -- Age at Entrance
2 -- Religious Affiliation
3 -- Marital Status
4 -- High School Grade Point Average
5 -- Mother's Occupation
6 -- Entrance Cohort
7 -- Fulltime/parttime Status
8 -- Major Field
9 -- Credits Earned in First Interval
10 -- Competence Level Units in First Interval
11 -- Prior Six Characteristics Rating

field. The number of credits accumulated by second assessment was related to fulltime status, major field, and age; competence level units accumulated were related to fulltime status, major field, high school grade average, and mother's occupation. The partial correlations presented in Table 12 control for both these influences on semesters, credits, and competence level units, as well as background and program covariates of ICS performance scores. The same is true for Six Performance Characteristics ratings, which were related to high school grade average and entrance cohort.

The significant correlation between number of semesters and the In-Basket score indicated that students who attended more semesters tended to receive lower scores. Most students attended at least the four regular semesters falling between entrance and second assessment, so the variability in number of semesters was limited to attendance during summer sessions. Thus, the data indicated that students who attended summer sessions tended to receive lower In-Basket exercise scores. There was a small but significant correlation between In-Basket scores and the first Six Characteristics rating. This correlation may be regarded as reciprocally supportive of the validity of the In-Basket exercise score as a measure of performance, and of the Six Characteristics ratings as a faculty rating of student performance.

The significant correlation between number of credits accumulated and Group Discussion scores indicated that students who accumulated more credits tended to receive lower scores on the Group Discussion exercise, with the influences of major, fulltime status, marital status and age controlled. In the case of the negative correlation between semesters and In-Basket scores, we may speculate that the association arises from the fact that students needing basic skills education and students re-taking a failed course populate the summer sessions (though we haven't the necessary data to

verify this guess). The negative correlation between Group Discussion scores and number of credits accumulated is more puzzling.

The only significant partial correlations between ICS scores and later measures of college performance were with number of semesters between second and third assessment. The number of semesters in this interval ranged only between three and five, with nearly three-quarters of the students attending the normal four semesters. This limited range allows the interpretation that students who attended summer school between their third and fourth years received higher Group Discussion and In-Basket scores, a finding for which there is no obvious explanation.

The partial correlations in Table 12 indicate there is no single performance variable underlying the several indices of college performance. The correlations between Group Discussion exercise scores and credits and semesters provided evidence that Group Discussion scores were not without some reliability, but do not necessarily indicate validity for the intended purposes of the score as a performance measure.

Relationships Between ICS and Human Potential Measures

As stated earlier, the repeated administrations of the human potential measures at entrance, two years after entrance, and again three and one-half years after entrance, provided three sets of correlations with ICS scores in different temporal relationships. Tables 13, 14, and 15, present the partial correlations between ICS scores and human potential measures for each temporal relationship, respectively. In each case, background and program influences on ICS scores and human potential measures were controlled.

Table 13.

Partial Correlations of Integrated Competence Seminar Scores with Entrance Assessments on the Human Potential Measures

	Exercises			
	Group Discussion	In-Basket (2)	Oral Presentation (3,7,8)	Total Score (2,3,7)
Measures of Vocational, Educational and Personal Issues				
"Best Class" Essay (8)	-.015 (118)	.015 (115)	-.018 (116)	-.028 (114)
"Decision" Essay (1,2,5,8)	.020 (95)	.005 (95)	.063 (93)	.065 (94)
"Career" Essay (1,4,8)	.072 (92)	.068 (89)	.030 (90)	.071 (88)
Sentence Completion Test (8)	-.139 (118)	.258**(115)	.102 (116)	.083 (114)
Defining Issues Test				
P% Score (1)	-.075 (119)	.175* (116)	.018 (117)	.065 (115)
D Score (1)	-.182*(119)	.147 (116)	.037 (117)	.020 (115)
Test of Cognitive Development (5,8)	-.084 (95)	.120 (93)	-.025 (93)	-.044 (92)
Picture Story Exercise				
Receptive (8)	-.001 (119)	.017 (116)	.046 (117)	.015 (115)
Autonomous	-.100 (122)	.067 (117)	-.087 (118)	-.056 (116)
Assertive (2,6)	-.135 (116)	-.109 (116)	-.156*(114)	-.179*(115)
Integrative (8,10)	-.068 (118)	.017 (116)	-.056 (117)	-.037 (115)
Self-Definition (8,10)	.011 (118)	-.046 (116)	.035 (117)	.008 (115)
Achievement Motive (8,9)	.034 (118)	.035 (115)	.141 (116)	.099 (114)
Affiliation Motive (8)	.000 (119)	.077 (116)	.179*(117)	.106 (115)
Power Motive	-.250**(122)	.027 (117)	-.228**(118)	-.197*(116)
Learning Style Inventory				
Concrete Experience (1)	.030 (123)	-.169*(120)	.001 (121)	-.056 (119)
Reflective Observation	-.067 (126)	-.057 (120)	-.082 (122)	-.122 (120)
Abstract Conceptualization (1)	-.049 (123)	.152 (120)	.009 (121)	.071 (119)
Active Experimentation	-.014 (126)	-.009 (120)	-.114 (122)	-.039 (120)
Abstract/Concrete Learning Orientation (1)	-.050 (123)	.197*(120)	.006 (121)	.080 (119)
Active/Reflective Learning Orientation	.035 (126)	.088 (121)	-.010 (122)	.055 (120)
Test of Thematic Analysis (8)	-.044 (121)	.216**(118)	-.046 (119)	.098 (117)

*p < .05
**p < .01

Table 13. continued

Critical Thinking Appraisal				
Inference	-.038 (117)	.206*(112)	.120 (113)	.119 (111)
Recognition	-.180 (117)	.073 (112)	.248**(113)	.024 (111)
Deduction (3,5)	-.223*(76)	.168 (75)	.224*(74)	.045 (74)

*p < .05
**p < .01

Note. Numbers in parentheses in cells are sample sizes. Numbers in parentheses in headings are covariates as follows:

1 -- Age
2 -- Religious Affiliation
3 -- Father's Education
4 -- Mother's Occupation
5 -- High School Grade Point Average
6 -- Prior College Experience
7 -- Marital Status
8 -- Entrance Cohort
9 -- Residence
10 -- Fulltime/parttime
11 -- Major

Table 14.

Partial Correlations of Integrated Competence Seminar Scores with Concurrent Change on the Human Potential Measures

	Exercises			
	Group Discussion	In-Basket (2)	Oral Presentation (3,7,8)	Total Score (2,3,7)
Measures of Vocational, Educational and Personal Issues				
"Best Class" Essay	-.039 (118)	-.087 (114)	-.080 (115)	-.100 (113)
"Decision" Essay (1)	.127 (118)	-.173*(115)	.055 (116)	.011 (114)
"Career" Essay	.156 (103)	.022 (100)	.020 (101)	.104 (99)
Sentence Completion Test (8)	-.001 (117)	.045 (114)	.062 (115)	.025 (113)
Defining Issues Test				
P% Score	-.075 (119)	.120 (116)	.165 (117)	.050 (115)
D Score	.008 (119)	.099 (116)	.108 (117)	.059 (115)
Test of Cognitive Development (8)	.042 (116)	.159*(113)	.146 (114)	.114 (112)
Picture Story Exercise				
Receptive (8)	.008 (118)	.060 (115)	.053 (116)	.048 (114)
Autonomous	.049 (119)	.061 (116)	-.041 (117)	.066 (115)
Assertive	-.002 (119)	.113 (116)	.016 (117)	.053 (115)
Integrative	-.197*(119)	.031 (116)	-.009 (117)	-.116 (115)
Self-Definition (10)	-.024 (118)	.088 (116)	-.023 (117)	.017 (115)
Achievement Motive	-.036 (119)	.112 (116)	.153*(117)	.081 (115)
Affiliation Motive (8)	.005 (118)	.041 (115)	.077 (116)	.061 (114)
Power Motive	.060 (119)	.109 (116)	-.014 (117)	.068 (115)
Learning Style Inventory				
Concrete Experience	-.176*(123)	.064 (120)	-.104 (121)	-.089 (119)
Reflective Observation	-.122 (123)	.028 (120)	-.159*(121)	-.155*(119)
Abstract Conceptualization	.080 (123)	-.110 (120)	.068 (121)	.026 (119)
Active Experimentation	.142 (123)	-.052 (120)	.194*(121)	.130 (119)
Abstract/Concrete Learning Orientation	.144 (123)	-.108 (120)	.096 (121)	.060 (119)
Active/Reflective Learning Orientation	.147 (123)	.015 (120)	.196*(121)	.153 (119)
Test of Thematic Analysis (8)	-.040 (120)	.088 (117)	-.031 (118)	.013 (116)

*p < .05

Table 14. continued

Critical Thinking Appraisal				
Inference	-.091 (114)	.144 (111)	.020 (112)	.006 (110)
Recognition	.086 (114)	.141 (111)	.034 (112)	.103 (110)
Deduction (5)	-.038 (92)	.102 (90)	.139 (90)	.046 (89)

See note on bottom of Table 13 for legend
*p < .05
**p < .01
***p < .001

Note. All human potential measures at second assessment were controlled for differences at entrance.

Table 15.

Partial Correlations of Integrated Competence Seminar Scores with Later Change on the Human Potential Measures

	Exercises			
	Group Discussion	In-Basket (2)	Oral Presentation (3,7,8)	Total Score (2,3,7)
Measure of Vocational, Educational and Personal Issues				
"Best Class" Essay	-.154*(119)	.012 (115)	-.028 (116)	-.097 (114)
"Decision" Essay (2)	-.152*(116)	.062 (113)	.007 (112)	-.046 (113)
"Career" Essay (8)	-.065 (117)	.005 (114)	-.060 (115)	-.049 (113)
Sentence Completion Test	.011 (118)	-.005 (115)	-.106 (116)	-.057 (114)
Defining Issues Test				
P% Score	.077 (118)	.113 (115)	.117 (116)	.122 (114)
D Score (1)	-.042 (118)	-.048 (115)	.006 (116)	-.030 (114)
Test of Cognitive Development	.070 (117)	.205*(114)	.037 (115)	.146 (113)
Picture Story Exercise				
Receptive	.034 (118)	.055 (115)	-.102 (116)	.017 (114)
Autonomous	-.026 (118)	.083 (115)	.242**(116)	.125 (114)
Assertive	-.003 (118)	.042 (115)	.114 (116)	.071 (114)
Integrative	.057 (118)	-.090 (115)	-.152*(116)	-.080 (114)
Self-Definition (8,10)	.046 (117)	.072 (115)	-.007 (116)	.053 (114)
Achievement Motive	.008 (119)	-.029 (116)	-.001 (117)	-.011 (115)
Affiliation Motive (8)	.015 (118)	-.024 (115)	-.045 (116)	-.023 (114)
Power Motive	.012 (119)	.177*(116)	.115 (117)	.140 (115)
Learning Style Inventory				
Concrete Experience	-.052 (123)	-.111 (120)	-.013 (121)	-.076 (119)
Reflective Observation	-.100 (123)	-.129 (121)	-.103 (121)	-.078 (119)
Abstract Conceptualization	-.059 (123)	.075 (120)	-.081 (121)	-.028 (119)
Active Experimentation	.012 (123)	.013 (121)	.120 (122)	.040 (119)
Abstract/Concrete Learning Orientation	-.009 (123)	.106 (120)	-.043 (121)	.023 (119)
Active/Reflective Learning Orientation	.065 (123)	-.049 (120)	.119 (121)	.063 (119)
Test of Thematic Analysis	.030 (121)	.109 (118)	.200*(119)	.119 (117)

*p < .05
**p < .01

Table 15. continued

Critical Thinking Appraisal				
Inference	.012 (114)	.086 (111)	.089 (112)	.068 (110)
Recognition	-.105 (114)	-.086 (111)	-.126 (112)	-.141 (110)
Deduction	-.070 (114)	.083 (111)	.129 (112)	.056 (110)

See note on bottom of Table 13 for legend.

*p < .05
**p < .01
***p < .001

Note. All human potential measures at third assessment were controlled for differences at second assessment.

Tables 14 and 15 include controls for preceding assessments on the same instruments, so that, in effect, the partial correlation with ICS scores represents the association of ICS scores with change on the human potential measures unexplained by the preceding assessment and any background or program covariates of change.

Table 13 presents the partial correlations between ICS scores and entrance assessments on the human potential measures. Variability in performance on the Integrated Competence Seminar or the human potential measures which could be attributed to background and program differences was controlled in all cases. Entrance cohort was a significant correlate of entrance assessments in the majority of cases shown, underscoring the importance of using multiple cohorts in program evaluation and studies of student change in college. Age was a significant correlate on three human potential measures despite the small variability in age in the Weekday College samples: on our measure of Perry's scheme of intellectual and ethical development, Rest's objective measure of moral development, and the measure of learning styles developed by David Kolb.

There were few significant partial correlations between entrance assessments on the human potential measures and the ICS exercise scores or ICS total score. The few small but significant partial correlations provide strongest support for the validity of the In-Basket exercise as a generic measure of performance in college. There is less, but some supportive evidence for the validity of the Oral Presentation exercise, and counter-evidence for the validity of the Group Discussion exercise.

In-Basket exercise scores were correlated with entrance measures of ego development, moral development, learning style, and critical thinking skills. Oral Presentation exercise scores were correlated with entrance measures of critical thinking skills, but were also negatively correlated with measures of self-definition and power motivation. The latter relationships do not directly support or invalidate the Oral Presentation exercise score. Group Discussion exercise scores were negatively correlated with a measure of moral development, with power motivation, and with a measure of critical thinking. These relationships indicate that students who did less well on the test of deductive thinking ability, those with lower scores on one measure of principled thinking, and students with less motivation to influence others, received higher scores on the Group Discussion exercise. This counter-evidence was not equivocated by any significant positive relationships. The data suggest that some entrance measures predict performance on the In-Basket, and the Group Discussion.

Tables 14 and 15 describe relationships between ICS exercise scores and total score and change on the measures of human potential for two successive intervals of time. Table 14 presents partial correlations between ICS scores and change on the human potential measures between entrance and second assessment, two years after entrance. This period approximates the general education sequence for most students in the sample and is most related to the time of student performance on the ICS. Table 15 presents the partial correlations between ICS scores and change on the human potential measures between the second and third longitudinal assessments, a period approximating the pre-professional sequence for most students in the sample, and usually follows ICS performance.

There were few patterns of significant relationship evident, but as in the case of Table 13, what evidence there was provides strongest support for relationships between the In-Basket exercise and less support for the Oral Presentation exericse. The relationships between Group Discussion scores and change on the human potential measures were negative. Overall, the statistically significant relationships were small enough and few enough, given the number of correlations generated, that the weight of evidence does not support the relationships of the Integrated Competence Seminar performance scores to change on the Human Potential Measures.

Summary and Discussion

Preliminary studies on the Integrated Competence Seminar indicated adequate reliability in the scoring procedure, and there was evidence of discrimination in individual performance. The relationships examined here between the ICS scores and other measures of student performance do not contribute much support for relationships to at least one if not two of the exercise scores. In our judgment, the In-Basket exercise has been shown to be the most difficult and the most related to other measures, the Oral Presentation exercise next most difficult and next most related to other measures, and the Group Discussion exercise the least difficult and the least related to other generic measures of student performance in college.

There are some possible explanations for these results. One has to do with performance mode. Another has to do with where the instrument is in its development as a measure of integration of abilities achieved by the midpoint of college.

The In-Basket was most related to performance in the learning process and to change on the measures of human potential. It is also a measure calling for problem-solving and critical thinking, college abilities more traditional in content. While the performance mode is nontraditional, it differs in mode from the Oral Presentation and Group Discussion. In the In-Basket, a student performs with respect to a particular paper and pencil task. The student does take on a role, but that is also true of the other two ICS exercises. The Oral Presentation and Group Discussion involve a speech and participation in a group as performance modes. It could be that these more nontraditional modes elicit aspects of Communications and Social Interaction not tapped by the human potential measures.

On the other hand, the evidence for negative relationships between Group Discussion and the college performance and human potential measures could be evidence that the measure is ineffective and invalid.

In fact, even before these analyses were complete, the Assessment Committee, the group of faculty responsible for the design and validity of the ICS, and external assessors, had raised questions about the acceptability and adequacy of the instrument stimulus, particularly related to the performance elicited by Group Discussion. They felt the group task did not elicit identification of complex issues or task development and solution as intended. An instrument revision process was undertaken after the data examined here were collected, so that a different instrument stimulus is in use now. While it would not be possible to directly follow up the implications of this analysis because of the change in instruments, the results have general implications for future development and research, particularly on the use of the performance modes used in this study.

The findings of this research, in terms of the relative validity of the exercises, suggest continuing to use the In-Basket, and to strive for better means to elicit the important abilities targeted in the Group Discussion and to further examine the Oral Presentation exercise. The Oral Presentation and Group Discussion exercises attempt to assess Communications and Social Interaction abilities that have not been traditional goals of teaching in liberal education, but are easily recognized as critically important abilities in the world of work. Further, the performance modes -- speaking and interacting -- are new assessment modes.

A related study of Social Interaction competence (Friedman, Mentkowski, Deutsch, Shovar and Allen, 1982) found that instructed and uninstructed students did not differ in the Group Discussion exercise. But when a student performance from a more heterogeneous group than were available for the current study was subjected to a discriminant analysis, instructed and uninstructed groups did separate on Social Interaction ability.

Thus, we may have the nucleus of an ability measure in Group Discussion as a performance mode. Only further research on the new ICS instrument stimulus will help answer this question. In sum, the study lands some support for the In-Basket, and suggests further research with Oral Presentation and Group Discussion exercises. While the current measure provided reliable results for diagnostic purposes, the new Integrated Competence Seminar needs to be researched further to continue its development as an external criterion measure.

References

Assessment Committee/Office of Research and Evaluation. Validating assessment techniques in an outcome-centered liberal arts curriculum: Six Performance Characteristics Rating. Milwaukee, WI: Alverno Productions, 1983.

Atkinson, J. W. Thematic apperceptive measurement of motives within the context of a theory of motivation. In J. W. Atkinson (Ed.), Motives in fantasy, action, and society. Princeton, NJ: Van Nostrand, 1958.

Friedman, M., Mentkowski, M., Deutsch, B., Shovar, M. N., & Allen, Z. Validating assessment techniques in an outcome-centered liberal arts curriculum: Social Interaction Generic Instrument. Milwaukee, WI: Alverno Productions, 1982.

Knefelkamp, L. L. Developmental instruction: Fostering intellectual and personal growth in college students. Unpublished doctoral dissertation, University of Minnesota, 1974.

Kolb, D. The Learning Style Inventory. Boston, MA: McBer and Company, 1976.

Loevinger, J., Wessler, R., and Redmore, C. Measuring ego development. Vol. II. San Francisco: Jossey-Bass, 1970.

McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. The achievement motive. New York: Appleton-Century-Crofts, 1953.

Mentkowski, M. Creating a "mindset" for evaluating a liberal arts curriculum where "valuing" is a major outcome. In L. Kuhmerker, M. Mentkowski, & V. Lois Erickson (Eds.), Evaluating moral development: And evaluating educational programs that have a value dimension. Schenectady, New York: Character Research Press, 1980.

Mentkowski, M., & Doherty, A. Careering after college: Establishing the validity of abilities learned in college for later success (NIE-G-77-0058). Milwaukee, WI: Alverno College, 1977.

Mentkowski, M., & Doherty, A. Careering after college: Establishing the validity of abilities learned in college for later careering and professional performance. Final Report to the National Institute of Education. Milwaukee, WI: Alverno Productions, 1983.

Mentkowski, M., & Strait, M. A longitudinal study of student change in cognitive development and generic abilities in an outcome-centered liberal arts curriculum. Milwaukee, WI: Alverno Productions, 1983.

Renner, J., Fuller, R., Lockhead, J., Johns, J., Tomlinson-Keasey, C., & Campbell, T. Test of Cognitive Development. Norman, OK: University of Oklahoma, 1976.

Rest, J. Development in judging moral issues. Minneapolis, MN: University of Minnesota Press, 1979.

Stewart, A. J. Analysis of Argument: An empirically-derived measure of intellectual flexibility. Boston: McBer, 1977. (a)

Stewart, A. J. Scoring manual for stages of psychological adaptation to the environment. Unpublished manuscript, Department of Psychology, Boston University, 1977. (b)

Stewart, A. J., & Winter, D. G. Self-definition and social definition in women. Journal of Personality, 1974, 42(2), 238-259.

Watson, G., & Glaser, E. Critical Thinking Appraisal. New York: Harcourt, Brace, Jovanovich, Inc., 1964.

Widick, C. An evaluation of developmental instruction in a university setting. Unpublished doctoral dissertation, University of Minnesota, 1975.

Winter, D. G. The power motive. New York: Free Press, 1973.

Winter, D. Test of Thematic Analysis. Boston, MA: McBer and Company, 1976.

APPENDIX I

Name_____ Score_____

A

INDEPENDENT JUDGEMENT

As Often AS Possible	Frequently	At Least Once	Not At All
A I	F I	O I	N I

GROUP DISCUSSION

INITIATION

2. Initiative - Active efforts to influence events rather than passive acceptance A F O N

INFLUENCE

3. Persistence - tendency to stay with a problem or line of thought until the matter is settled A F O N

4. Impact - Ability to create : Good Impression A F O N

Attention and respect A F O N

Show an air of confidence A F O N

5. Social effectiveness - Ability to bring a group to accomplishment of its goal by directing and facilitating actions A F O N

Response

6. Listening - ability to understand what was significant in oral communication of others. A F O N

7. Flexibility - Intellectual , emotional, social - ability to modify behavioral style to reach a goal. A F O N

8. Adaptability - to new situations. Ability to assume the role called for by a given situation. A F O N

PROBLEM ANALYSIS

9. Organization - effectiveness in seeking out pertinent data and organizing it to arrive at a credible conclusion (Research, budget, ABC plan) A F O N

10.Raising of questions A F O N

Decision Making

11. Consideration and evelauation of alternatives desisions, conclusions behavior A F O N

12. Decisiveness - Readiness to make decisions or to render judgement A F O N

Total _____

% _____

NAME_____ IN - BASKET

Setting Priorities 1 0

Complete Exercise 1 2 3 4 5 6

INITIATION

1. Clear convincing writing which does not distract the reader by technical errors. A F O N
2. Initiative - Active efforts to influence events rather than passive acceptance. A F O N

INFLUENCE

3. Impact - ability to create : Good impression A F O N
   Command attention and respect A F O N
   Show an air of confidence A F O N

RESPONSE

4. Flexibility - Intellectual, emotional, social - ability to modify behavioral style to reach a goal A F O N
5. Adaptability - to new situations A F O N
6. Reading - ability to understand what was significant in each item. A F O N

PROBLRM ANALYSIS

7. Concern for significant details A F O N
8. Making predictions based on pertinent data ability and willingness to project future implications of present situations and or decisions. A F O N

DECISION MAKING

9. Consideration and evaluation of alternatives: Decisions, conclusions behavior A F O N
10. Decisiveness - readiness to make decisions or to render judgement A F O N

Total _____

% _____

ORAL PRESENTATION

NAME _____

INITIATION

1. Oral expression - clear understandable communication of message A F O N

INFLUENCE

2. Impact - ability to create : A good impression A F O N
   Command attention and respect A F O N
   Show an air of confidence A F O N

RESPONSE

3. Reading - ability to understand what was significant in source materials A F O N
4. Adaptability to new situations - ability to assume the role called for by a given situation A F O N

PROBLEM ANALYSIS

5. Organization - effectiveness in seeking out pertinent data and organizing it to arrive at a credible conclusion A F O N
6. Concern - for significant detail - ability to cite and utilize whatever can provide clarification or evidence (e.g. examples, observations, facts distinctions relationship). A F O N

DECISION MAKING

7. Consideration and evaluation of alternatives : Decisions, Conclusions, Behavior. A F O N
8. Decisiveness - readiness to make decisions or to render judgement A F O N
9. Value clarification - ability to identify values in decision making A F O N

Total _____

% _____

3401 South 39th Street / Milwaukee, WI 53215